Identification of molecules that interact with the adaptor protein Dof

Inaugural-Dissertation

zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Alysia Battersby
aus Kapstadt, Südafrika

Köln 2001

Berichterstatter: Prof. Dr. Maria Leptin
Prof. Dr. Karin Schnetz

Tag der Disputation: 3 Juli 2001

Table of Contents

1. INTRODUCTION

A key requirement in metazoan development is for cells to communicate with each other. As a zygote gives rise to a body of cells that will ultimately make the adult organism, groups of cells will acquire distinct fates in order to contribute to distinct organs or tissues. For the various aspects of morphogenesis to occur, cells must remain in constant communication with each other, receiving or sending instructive, inhibitory or permissive signals. Given the diversity of biological responses found in cells during the course of development – proliferation, differentiation, migration, cell metabolism and cell survival to name just a few – it is surprising that the signal transduction pathways that elicit unique responses in different target tissues often make use of the same molecules. Thus, the diversity of signals is not paralled by a diversity of molecules. Furthermore, pathways are not linear and extensive crosstalk amongst pathways exists. One of the challenges in understanding signal transduction is to define which components and what combinations of crosstalk make each biological response unique.

The embryo of the fruitfly *Drosophila melanogaster* provides a well developed experimental system to investigate the nature of a signaling event. In addition to its being amenable to genetic dissection, a handful of signaling cascades that have been well characterized allow us to address the specificity of a signal in relation to its unique biological response. This is particularly clear in the case of Receptor Tyrosine Kinase (RTK) signaling. Several RTK cascades have been characterized and as mentioned above, despite the diverse biological responses, many of their components are shared. Thus, all RTKs can activate the Mitogen Activated Protein Kinase (MAPK) cascade via the small GTP-binding protein Ras. The ultimate response elicited by each RTK is, however, unique. It therefore remains to be determined what features of signaling are combined to achieve this effect.

The fibroblast growth factor (FGF) signaling pathways in *Drosophila* provide a suitable system to analyze two well described effects of RTK signaling via the Ras-MAPK signaling pathway, namely differentiation and cell movement. *Drosophila* is equipped with two FGF receptors (FGFRs) called Heartless and Breathless. Heartless (DFGFR-2) is required early in development for the cells of the mesoderm to spread along the lateral ectoderm and later for proper differentiation of heart precursors and midline glia cells (Shishido et al. 1997; Michelson et al. 1998). Breathless (DFGFR-1) plays multiple roles in the formation of the respiratory system. It is initially used to establish the tracheal branch network and later employed to remodel tracheal branches in response to changes in oxygen requirement (Jarecki et al. 1999; Klämbt et al. 1992; Reichman-Fried et al. 1994). A subtle role of Breathless in the migration of a subset of glia cells has also been described (Klämbt et al. 1992). While both FGF pathways have been shown to signal through Ras, the pathways to their effects are by no means linear and branch points remain to be elucidated. Downstream of FGFR (Dof), a cytoplasmic molecule essential exclusively for FGF signal transduction (Imam et al. 1999; Michelson et al. 1998; Vincent et al. 1998), lies upstream of the Ras-MAPK pathway and provides an entry point into this investigation. Dof provides us with a signaling component that uses common cascades but confers specificity to one type of RTK signal, namely that of the FGFR. The way this large molecule functions is unknown, and one way to gain insight

into its function along with defining potential branch points of the pathway is to identify interacting partners that are part of the signaling cascade. In this work, a yeast two hybrid screen was conducted to identify such partners.

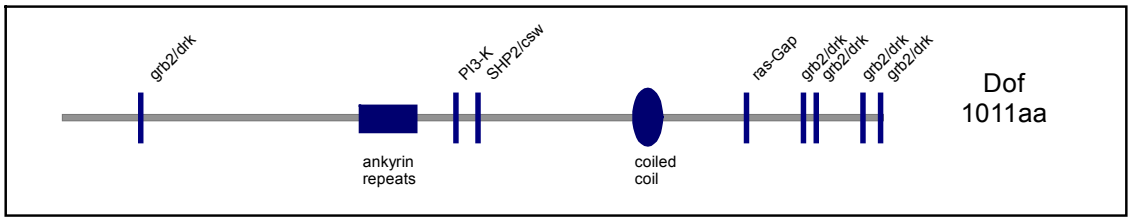


Figure 1.1 This is a figure of the predicted motifs and structures in the Dof protein (transcript II) adapted from Vincent et al. (1998). The ankyrin repeat is marked by a square and the coiled coil domain by an oval. Vertical bars indicate the position of tyrosines which, when phosphorylated, might be binding sites for the proteins indicated above the diagram.

Dof can be defined as an adaptor molecule because it links upstream and downstream molecules of a signaling cascade. The Dof protein contains an ankyrin repeat and a coiled-coil domain, two features which confer Dof with an homology to the vertebrate tyrosine kinase substrate BCAP (Okada et al. 2000). Dof also contains several tyrosines within environments that become putative binding sites if the tyrosine is phosphorylated. The motifs are shown in Figure 1.1. However, although Dof is tyrosine phosphorylated in the presence of the activated FGFR Breathless in tissue culture, none of the putative tyrosines shown in Figure 1.1 is essential on its own for Dof to function (R. Wilson personal communication). Thus, apart from the fact that Dof is required for the transduction of signals by FGFRs, and that an artificially activated Ras molecule can more or less compensate for its absence, not much else is known about how Dof signals and can only be guessed by looking at what is already known about a cell that engages in FGF signaling and what goes wrong when this does not occur. The introduction to this thesis attempts to assemble information about signaling pathways and molecules that have been characterized in more detail (many reviews are available on this vast topic), and hopes to refine the expectations to be put on a molecule whose lack of functional structural motifs make it difficult to place in any of the known signaling categories.

1.1 Signal Transduction via Receptor Tyrosine Kinases

Receptor Tyrosine Kinase activation

Receptor Tyrosine Kinases constitute a large group of membrane spanning cell surface receptors that activate cytoplasmic signaling cascades and elicit a diverse range of responses in target tissues (reviewed in Schlessinger 2000 and references therein). As their name suggests, RTKs have intrinsic protein tyrosine kinase activity. The tyrosine kinase is situated in the cytoplasmic domain. When a ligand binds to the extracellular domain of the receptor, it induces receptor dimerization, bringing the protein tyrosine kinases of RTK monomers close enough to phosphorylate each other at tyrosine residues in the activation loop of the catalytic domain. Phosphorylation causes the activation loop to adopt an open configuration, permitting access of the kinase to ATP and substrates.

The active, autophosphorylated receptor dimer transmits the signal in two ways: First, its active kinase can phosphorylate downstream signaling proteins. Second, the active receptor can phosphorylate further tyrosines on its cytoplasmic domain (mostly in the noncatalytic regions of the receptor molecule) which thereby become docking sites for adaptor molecules, which it can phosphorylate, too. Other adaptor molecules might be constitutively bound to the cytoplasmic domain, and become phosphorylated upon kinase activation.

Signaling cassettes activated by Receptor Tyrosine Kinases

Receptor tyrosine kinases activate a number of signaling cascades. As Figure 1.2 illustrates, receptor tyrosine kinases activate the Ras MAPK cascade, phospholipase Cγ (PLCγ), Phospholipid Kinase (PI_3K) and the JAK/STAT pathway and thus activate effector molecules in the nucleus or the cytoplasm. The arrows in this flow chart emphasize that the signaling cassettes activated by the receptor tyrosine kinase engage in cross talk, thus creating a complex network of stimulatory and inhibitory signals. The activation of the Ras MAPK cascade will be described below in more detail, since it is essential for transduction of the FGF signal in *Drosophila*. However, this figure is intended to expose the scope of possible signaling pathways that can be used by a receptor tyrosine kinase, given the appropriate intracellular domain and the matching intracellular proteins to respond to its activated state.

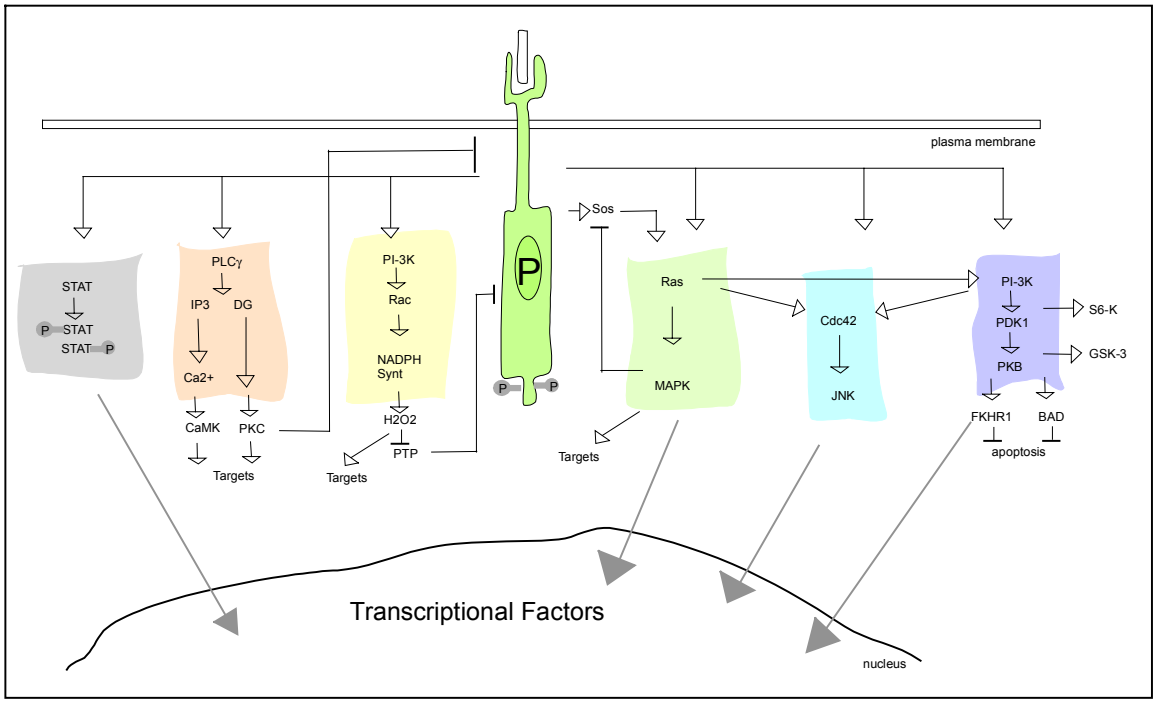


Figure 1.2 This figure has been adapted from Schlessinger (2000) and illustrates the various signaling cassettes (not all components are included) that have been shown to be activated by RTKs. The autophosphorylated receptor is shown in green and the various signaling casettes are depicted in shaded boxes.

Receptor Tyrosine Kinase activation of the Ras MAPK cascade

Many cell surface receptors, including all RTKs, stimulate the exchange of GTP for GDP on the small G protein Ras. Ras is activated by the guanine nucleotide exchange factor Son of sevenless (Sos), which is constitutively complexed with the adaptor protein Grb2 via its SH3 domain. Depending on the receptor, the Sos/Grb2 complex is recruited to an activated RTK directly or via intermediate molecules. The SH2 domain of Grb2 can bind to specific phospohotyrosine sites of the receptor and thereby position Sos at the plasma membrane, where it is brought into the vicinity of Ras. Ras is anchored at the inner face of the plasma membrane via C-terminal lipid modification (reviewed in Hunter 2000). Alternatively, Grb2/Sos can bind to the adaptor protein Shc, which forms a complex with many receptors through its phosphotyrosine (PTB) domain. Another means for activation of the Ras MAPK pathway is via the binding of Grb2/Sos complexes to membrane linked docking proteins such as insulin receptor substrate-1 (IRS1) or FRS2α, which become tyrosine phosphorylated in response to activation of certain RTKs.

Ras activates MAPK cascade and PI_3K

The active GTP-bound Ras interacts with several effector proteins. The best characterized effector proteins are Raf and PI(3)K (reviewed in Rommel and Hafen 1998). Raf is the protein kinase at the top of the MAPK cascade that leads to the activation of ERK (extracellular signal related kinases). The phospholipids generated by activated PI(3) kinases serve as second messengers and recently, distinct binding domains for 3-phosphoinositides in diverse range of target molecules have been identified (reviewed in Leevers et al. 1999). One of these bindings domain is the Pleckstrin homology (PH) domain, a module that can regulate a protein's membrane localization in response to signaling (see below).

Crosstalk between Ras and RhoGTPases

While many cellular events are affected by Ras GTPases, the Rho family of GTPases is considered to be the major orchestrator of actin mediated cytoskeletal changes, and has been shown to engage in crosstalk with Ras, namely the Rho family of GTPases (reviewed in Hall and Bar-Sagi, 2000). It has been shown that Cdc42, a member of the Rho family of GTPases, can be activated by Ras, thus providing a link between two major signaling pathways.

The highly conserved MAPK cascade

Mitogen-activated protein kinases (MAPK) are the convergence point of a multitude of signaling cascades and when activated influence a wide range of cellular processes such as metabolism, cell cycle, cell migration, cell shape, cell proliferation and differentiation (reviewed in Davis, 2000). Activated MAPKs can translocate into the nucleus and activate transcription factors and phosphorylate effector proteins in the cytoplasm, for example pp90 ribosomal S6 kinase (reviewed in Blenis, 1993).

MAPK is the serine-threonine protease at the bottom of a three-member protein kinase cascade. MAPK is activated by MAPK kinase (MAPKK) which in turn is activated by a MAPKK kinase. It is worth noting that the MAPKK kinases, that is, the kinases at the top of the modules, have different defined regulatory motifs that are not found in MAPKKs or MAPKs, including PH domains, SH3 binding domains, binding sites for GTP-binding proteins and phosphorylation sites for tyrosine and serine/threonine kinases – indicating that these can be differentially regulated by a variety of upstream inputs (Garrington and Johnson, 1999). The components of this cascade have been conserved from yeast to mammals and every eukaryotic organism has multiple MAPK pathways, which are largely separate from one another (reviewed inHunter, 2000). One factor that keeps the pathways separate is that the MAPK kinases phosphorylate their target MAPKs on conserved threonine and tyrosine residues which are separated by an intervening amino acid that is characteristic for each MAPK family.

Raf is the MAPKKK which activates MEK, a MAPKK. When MEK activates ERK, ERK translocates from the cytoplasm to the nucleus where it phosphorylates and activates transcription factors, for example ternary complex factor (reviewed in Karin and Hunter 1995).

1.2 Protein Modules

A salient feature of signaling molecules, including those that mediate signals from RTKs, is that their composition is modular (motifs and sequences are reviewed in Pawson 1995; Kuriyan and Cowburn 1997; Pawson and Scott, 1997). A modular protein is one which is made up of several domains, and in a way, each module is an entity on its own. Individual modules may have enzymatic properties or a domain that brings about interactions with other proteins, phospholipids or nucleic acids. By having several modules, catalysis is separated from target recognition, and target recognition modules can be combined in an as yet undefined number of ways. Modular signaling molecules create versatile combinations of input and output signals and it is easy to see how they contribute significantly to the branches within complex signaling networks.

There are two types of modules that bring about interactions with other molecules. Some modules, like Src homology-3 (SH3) domains and some phosphotyrosine binding (PTB) domains are the mere plugs and sockets that connect parts of a network together. Other modules, most prevalently the Src homology-2 (SH2) domain and many PTB domains act as sensors of transient states that are manifested in nothing more than a phosphorylated tyrosine in the interacting partner. In other words, SH2 domains and many PTB domains only bind to motifs in which the tyrosine is phosphorylated. Other modules, like Pleckstrin Homology (PH) domains regulate the protein's localization in the cell in response to signaling. This section introduces the modules have been found in signaling molecules that transmit signals from activated RTKs.

Domains that mediate binding to an activated RTK

The activated RTK oligomers can phosphorylate tyrosines on target proteins as well as tyrosines on itself. The tyrosines that are phosphorylated on the receptor serve as docking sites for SH2 or PTB containing proteins.

SH2 domains bind to amino acid sequences defined by 1 to 6 residues C-terminal to the phosphotyrosine moiety. One group of SH2 domain containing proteins contains intrinsic enzymatic activities that are activated by the RTK merely by it localizing them to the membrane via their SH2 domain (as in the case of PI3-K) or by a conformational change caused by SH2-domain-mediated binding to the receptor (as in the case of Src).

SH3 domains are small modules that bind to sequences containing proline residues arranged in characteristic PxxP motifs. Some RTK binding proteins contain only SH2 or SH3 domains and appear not to have any enzymatic activity. Their main role is to function as adaptors between molecules. For example, the adaptor protein Grb2 links a variety of surface receptors to the Ras/MAP kinase signaling cascade by interacting with activated RTKs via its SH2 domain and using its SH3 domain to recruit Sos close to its target protein Ras at the cell membrane.

PTB domains have been found to mediate binding between a RTK and a variety of signaling molecules. In fact, vertebrate FGFRs use PTB containing proteins, not SH2 domain proteins to recruit the Grb2/Sos complex to the membrane (Klint et al. 1995) Most, but not all PTB domains bind to a motif that requires a phosphotyrosine. For example, FRS2, a lipid-anchored docking protein (Kouhara et al. 1997) uses its PTB domain to constitutively bind to a the juxtamembrane region of FGFR1 (Xu et al. 1998, Ong et al. 2000), but tyrosine phosphorylation is not required, and the minimal sequence required for binding does not have any tyrosines (Ong et al. 2000). In contrast, the same PTB domain will bind to a classical NPQpY motif of the nerve growth factor (NGF) receptor TrkA (Ong et al, 2000). As can be seen by the example of FRS2, PTB domains are modules with diverse ligand-binding specificities.

Domains that mediate membrane localization

Localization of signaling molecules within the cell is an important feature of signal transduction. Not only can this provide directionality to a signal (as in chemotaxis), but some proteins can only function at certain locations in the cell. Most obviously, proteins that activate transcription have to be in the nucleus to function, and cytoplamic anchoring is used as a means of regulating their actions. For example, STAT (signal transducer and activator of transcription) resides in the cytoplasm and can only translocate to the nucleus once it is phosphorylated by Jak tyrosine kinase (reviewed in Darnell 1996; O'Shea 1997). Other signaling molecules have to be recruited to the plasma membrane in order to function. Some molecules, like Ras proteins are constitutively localized to the plasma membrane, whereas others localize to the membrane in response to a signal. This can be achieved by the modification of lipids in a region of the membrane or it can be achieved by creating binding sites in proteins that are already at the membrane, a feature that has been discussed above: SH2 and some PTB containing proteins translocate to the membrane by binding to phosphotyrosine sites in receptors or intermediate molecules.

Constitutive membrane localization

Lipid modifications of proteins can lead to their constitutive membrane localization. Core signaling molecules such as the Ras proteins are localized to the plasma membrane via farnesylation of the cysteine of the C-terminal CAAX motif in combination with palmitoylated or polybasic residues (reviewed in Magee and Marshal 1999). This membrane localization is crucial for Ras function. This is where it is activated by the Grb2-Sos nucleotide exchange factor and this is where it activates its downstream targets.

Transient membrane localization

PH (Pleckstrin homology) domains are found in a variety of signaling proteins, including certain guanine nucleotide exchange factors (GEFs), phospholipase-C (PLC)γ, and a large number of docking proteins (Schlessinger, 2000). Some PH domains have been shown to preferentially bind $PI(3,4)P_2$ and/or $PI(3,4,5)P_3$ over other inositol lipids (reviewed in Leevers 1999). 3-phosphoinositides are generated by PI3 kinases and different classes of PI3 kinases preferentially phosphorylate different lipids. When the PH domain of a protein binds to a certain inositol lipid, this not only allows its translocation to a specific membrane but might also induce a conformational change that allows its phophosphorylation by other kinases. For example, it has been suggested that the Akt-kinase PDK1 can only phosphorylate and activate Akt when its conformation has been changed by its binding to 3-phosphoinositides (Stokoe et al. 1997, Alessi et al. 1997).

PH modules are especially interesting because they have been implicated in directional movement. This was most clearly shown in *Dictyostelium discoideum* by tagging the PH domain containing proteins like CRAC (cytosolic regulator of adenylyl cyclase protein) and Akt with Green Fluorescent Protein (GFP) (Parent et al. 1998; Meili et al. 1999), and tracing their intracellular localization upon stimulation with chemoattractant. When stimulation occurs, CRAC-GFP, Akt-GFP as well as plain PH-GFP, transiently localize to the leading edge (reviewed in Parent and Devreotes, 1999). It is thought that the PH domain binds to Gβγ subunits or phosphorylated inositol lipids at the inner surface of the plasma membrane.

Docking proteins

Docking proteins are another category of protein that mediate membrane recruitment of signaling proteins stimulated by tyrosine phosphorylation of a receptor (reviewed in Schlessinger, 2000). All docking proteins contain in their N-termini a membrane targeting signal like a myristoyl anchor or their own transmembrane domain, but mostly a PH domain. The C-terminus has a large region that contains multiple binding sites for the SH2 domains of signaling proteins. Most docking proteins also contain domains like the PTB domain that are responsible for complex formation with a particular set of cell surface receptors. For example, the docking protein FRS2 binds specifically to FGFRs and NGRs (see above).

Scaffold proteins

Scaffold proteins are signal mediators because they bind to several proteins of a particular cascade, thereby providing a fast transmission of a signal and an insulation from cross-talk with other signaling molecules. Thus they provide another means of achieving specificity because they not only sequester proteins so that they do not interact with other proteins but can also regulate the activity of the proteins to which they bind, by for example bringing an enzyme into the vicinity of its substrate. A well studied example is Ste5, from the yeast *Saccharomyces cerevisiae* (reviewed in Garrington and Johnson, 1999). Three MAPK cascade modules exist in yeast and one of these modules transduces the mating response upon pheromone binding to a G-protein-coupled seven transmembrane receptor. The Gβγ subunit complex that is released from the activated receptor binds to the scaffolding protein Ste5, which in turn binds to all three components of the mating response MAPK module (Ste11, Ste7, Fus3), thus ensuring a rapid transmission of the signal and also an insulation from the activation of the other MAPK modules. One of the components, Ste11, is also used in another MAPK module, the high osmolarity response pathway, but there it is bound to another scaffold (Pbs2). Thus, for each cellular response there is a defined set of signaling ingredients, and scaffold proteins channel right ingredients into the appropriate direction.

1.3 Receptor Tyrosine Kinase signaling in *Drosophila*

Several RTKs and their signal transduction pathways have been characterized extensively in *Drosophila* (reviewed in MacDougall and Waterfield, 1996 and Sternberg and Alberola-Ila, 1998) and like all organisms, share multiple signaling components. Two of these RTKs are employed for very specific events: Sevenless is involved in photoreceptor specialization, and Torso in terminal structure differentiation. The *Drosophila* epidermal growth factor receptor (DER) is involved in multiple cellular events, including dorsal-ventral axis specification and wing-vein formation. As described earlier on, the two FGFRs in Drosophila are required for several processes, including mesoderm and heart development (Heartless) and tracheal outgrowth (Breathless).

All RTKs that have been identified in *Drosophila* can signal via the Ras-MAPK cascade (reviewed in Sternberg and Alberola-Ila, 1998). While the core of the Ras-MAPK cascade is used by all RTKs, each receptor can activate Ras1 via a different set of molecules. Thus, activated Torso, Sevenless and EGFR can activate Ras1 either directly by recruiting (and phosphorylating) grb2/drk-sos to its phosphotyrosines or via adaptor molecules such as Dos (Herbst et al. 1996, Raabe et al. 1996) and Shp2/Csw (Perkins et al. 1992). Recently, two intermediate molecules have been identified that selectively transduce a signal from specific RTKs. The adaptor/scaffolding molecule Shc has been shown to link Torso and EGFR, but not Sev, to the Ras MAPK pathway (Luschnig 2000). Similarly, Dof has been shown to link both FGFRs to the Ras MAPK pathway, but does not function downstream of Torso, Sevenless or DER (Vincent et al. 1998; Michelson et al. 1998; Imam et al. 1999).

A good example of the branching of a signal comes from double-mutant analysis of the Torso pathway. Luschnig et al. (2000) showed that the signal splits into at least three parallel branches, represented by Dos, Drk and Shc. But the components might

cross-talk, since Dos contains putative binding sites for the SH2 domains of Drk and Shc (Raabe et al. 1996), although the Drk/Grb2 SH2 domain binding sequence has been shown not to be essential in rescue experiments (Allard et al. 1998).

1.4 FGF Signaling in *Drosophila*

As described at the beginning of this Introduction, the two FGFRs in *Drosophila*, Heartless and Breathless are required for morphogenetic processes involving differentiation and cell movement. They elicit these diverse cellular responses by activating the Ras MAPK cascade, a processes which is mediated by the cytoplasmic protein Dof. Whether the FGFRs activate other signaling pathways, and whether Dof is required for the transduction of these, is unknown. In any case, it is unlikely that the diverse cellular responses elicited by FGF signal transduction, are performed by a single, linear pathway. The following section aims to portray the two best characterized cellular processes in *Drosophila* that employ FGF signaling, mesoderm development and tracheal branch outgrowth.

Early mesoderm development

The mesoderm germ layer of the *Drosophila* embryo gives rise to somatic and visceral musculature, the fat body and heart cell precursors (Bate 1993). The schematic in Figure 1.3A is a cross section of a *Drosophila* embryo whose presumptive mesoderm has just invaginated into the interior of the embryo and whose cells are just about to spread along the lateral ectoderm into a single layered sheet of cells (Figure 1.3B). Depending on their position along the dorso-ventral and anterior posterior axis, the cells receive instructions on their fate.

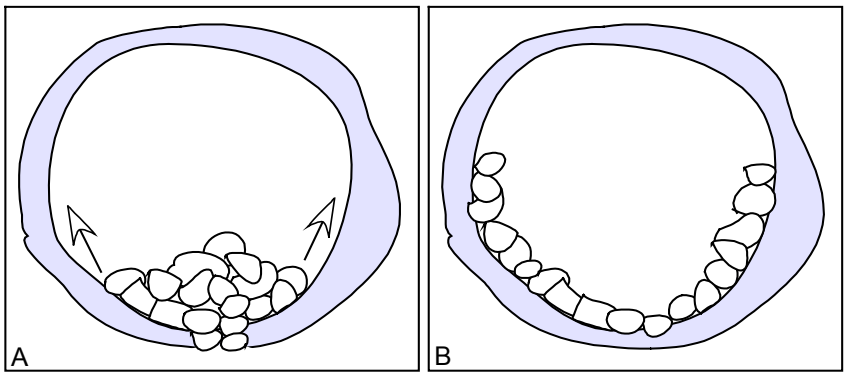


Figure 1.3 A. This is a simplified sketch of a cross section of an embryo whose mesoderm anlage has just invaginated and has undergone an epithelial to mesenchyme transition. Dorsal is to the top and ventral is to the bottom. The ectodermal germ layer is shown in blue and arrows indicate the direction in which the cells will spread. **B.** This sketch depicts the mesodermal cell layer after it has spread along the lateral ectoderm.

FGF signaling, via the FGFR Heartless is required for the mesoderm to spread and later for heart cell precursors to differentiate. The ligand which activates Heartless is unknown and it is unclear what directional cues generate the spreading movement. The

manner in which the cells spread has implications for the mechanisms used by the cells to reach their final positions along the ectoderm (Wilson and Leptin 2000). Experimental evidence supports the option in which the flattened layer of mesodermal cells (Figure 1.3B) is created by cells seeking to maximize their contact with the underlying ectoderm from which an instructive signal would be emanating. An interesting possibility for cells to reach out to the ectoderm is by using filopodial extensions to find and contact the ectoderm. Such filopodial extensions have been observed in imaginal disc cells, which send out long filopodial extensions towards sources of FGF (Ramirez–Weber and Kornberg, 1999)

Tracheal branch outgrowth

The larval tracheal system is made up of epithelial cells that form a hollow tubular network that delivers oxygen to target tissue via passive diffusion. The tracheal system develops from ten metameric tracheal placodes on each side of the embryo which invaginate to form tracheal pits. It is from these sac-like structures that branching morphogenesis occurs. The branched network has been described to form in 3 stages, each stage employing distinct cellular and molecular mechanisms to create each successive branch (Samakovlis et al. 1996) The branches occurring at the first, second and third stage are called the primary, secondary and terminal branches respectively, and these stages are depicted in the schematic below (Figure 1.4).

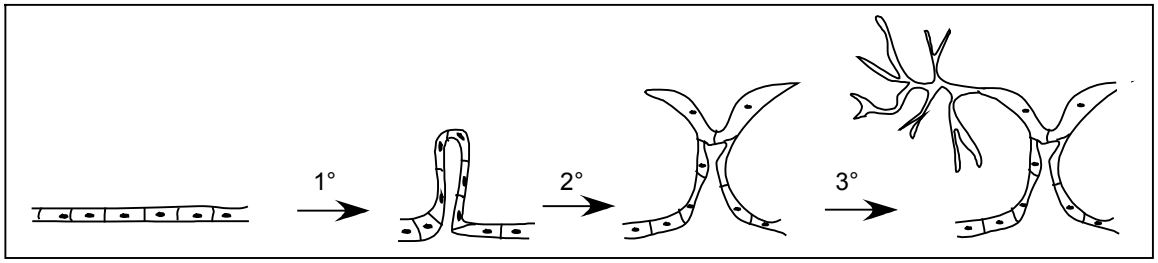


Figure 1.4 This figure is a schematic of the three stages of tracheal branching that occur from a portion of an epithelial sac. Picture adapted from Metzger et al. 1999.

The first two stages of tracheal branch outgrowth involve a multitude of stereotypic genetic programs whereas the third stage, during which terminal branches reach their targets, is determined by the oxygen requirement of target cells (Samaklovis et al. 1996). The chemoattractant Branchless is required at all stages of branching morphogenesis. The gene *branchless* encodes the FGF ligand for the Breathless receptor. Branching starts when primary and secondary branches respond to the Branchless, which is secreted by a stereotypic pattern of ectodermal and mesodermal cells that define the direction of branch outgrowth (Sutherland et al. 1996). The terminal branches are made out of cytoplasmic extensions that respond to the the oxygen requirement of target cells. In larvae, Branchless has been shown to mediate terminal outgrowth to oxygen deprived target tissue (Jarecki et al. 1999).

During tracheal branch outgrowth, the FGF signaling pathway must act at different levels. It is required for sheets of cells to move toward a morphogen (primary branches) and also for subcellular portions of the plasma membrane to make extensions

toward oxygen deprived tissue (terminal branches). If such distinct morphogenetic events are to be orchestrated by the FGF signaling pathway, its downstream components are surely intertwined with a vast number of other cellular events and signals whose contributions to the final tracheal tree remain to be elucidated.

1.5 Specific Aims of this thesis

Given the different cellular responses that are triggered by the FGF signal transduction pathway in Drosophila, an interesting question that arises is how one signal can generate different responses. One way of answering this question is to look for the components of a signaling system and determine how they interact with each other to bring about a distinct cellular response. The adaptor molecule Dof is an essential component of the FGF signaling system, and because it is used specifically by FGF receptors, makes for an interesting component of the pathway. In this work, a yeast two-hybrid screen was conducted to identify interacting partners of Dof.

2. RESULTS

In this project, a yeast two-hybrid screen was conducted to identify interaction partners of Dof. The Gal4 DNA-binding domain fused to a C-terminal deletion construct of Dof was used to screen an embryonic *Drosophila* cDNA library for potential binding partners of Dof. Two groups of interacting partners were isolated: those that activated the stringent reporter Ade2 and those that only activated the less stringent reporter His3. Three Ade2 positive candidates proved to be particularly interesting, namely the FGFR Heartless, the SUMOlating enzyme Ubc9, and Dof itself. These three candidates will be presented in detail in Chapters 2.3, 2.4 and 2.5. The final Chapter, Chapter 2.6 summarizes all the other candidates and serves primarily to document the groundwork that was made for yeast two-hybrid interacting partners whose involvement in Dof function remains unknown.

2.1 Description of components used in the yeast two-hybrid screen

The yeast two-hybrid assay

The yeast two-hybrid system devised by Fields and Song (1989) was used to provide a physical environment in which to detect potential interactions involving Dof. The yeast two-hybrid assay takes advantage of the fact that domains of transcription factors are modular. Thus, many transcription factors can be separated into at least two distinct functional domains, the activation domain and the DNA-binding domain (Keegan et al. 1986, Hope et al. 1986). These two domains can constitute a functional transcriptional activator even if they are not covalently attached to each other. An interaction between two proteins can be assayed by fusing one to the DNA-binding domain and the other to the DNA-activation domain (Fields and Song 1989). Upon interaction of the two proteins, a hybrid transcription factor is reconstituted and expression of a reporter gene linked to the upstream activating sequence is activated. The yeast two-hybrid system can thus be used to clone cDNAs encoding interaction partners of a known protein by screening an activation-tagged cDNA expression library. In the literature, hybrid proteins might be referred to as bait, prey, or activation tagged. These are terms used to refer to the various hybrids: proteins fused to the DNA-binding domain are termed “bait” and those fused to the activation domain are “activation tagged” and when part of a cDNA library, collectively called the “prey”.

Components of the Yeast Two-Hybrid system

To screen for cDNAs encoding proteins able to interact with Dof in the yeast two-hybrid system, the general scheme outlined in Figure 2.1 was employed.

Figure 2.1 This figure is an outline of the selection strategy used to isolate interaction partners of Dof in the yeast two-hybrid system. On the left, the bait construct DB-DofΔC does not activate the reporters when expressed on its own. Once co-expressed with another protein that is fused to the activation domain ("AD-candidate") activation of the reporters will occur if the two fusion proteins interact.

Yeast Strains

The host yeast strain used in this study was PJ69-4A (James et al. 1996). Like all strains engineered for use in the yeast two-hybrid system, PJ69-4A contains inducible promoter elements fused to reporter genes, thus providing a detectable response to protein-protein interactions. PJ69-4A carries three chromosomally located reporter genes, each under the control of a different Gal4 responsive promoter. Although all are induced to high levels by the transcription factor Gal4, they share a minimum of sequence identity. This reduces the number of promoter-specific false positives. When this strain is used in a yeast two-hybrid screen, a true interaction should activate all three markers. In practice, the His3 nutritional marker and the *Escherichia coli* lacZ gene are activated more readily that the Ade2 nutritional marker. This means that when using Ade2 to select for interactions, one should be aware of the fact that weak interactions might be overlooked (Phillip James, personal communication; see also Discussion).

Expression vectors

Expression vectors used in the yeast two-hybrid screen

The two fusion proteins used in a yeast two-hybrid assay are generated by two different cloning vectors. For the yeast two-hybrid screen with Dof, the bait vector generates a hybrid protein between the DNA-binding domain of Gal4 (pGBD-c(x), James et al. 1996) and a portion of the Dof protein. The prey vector (pACT2) contains sequences for the Gal4 activation region II (Durfee et al. 1993) fused to a cDNA library generated from 0-18hr *Drosophila* embryos.

Other expression vectors

Besides pGBD-c(x) and pACT2, other expression vectors were used to assay various yeast two-hybrid interactions. The reader is referred to Appendix I for a documentation of the different interaction levels observed by different combinations of expression vectors. All vectors contain a truncated version of the Adh1 promoter sequence, providing for constitutive expression of the recombinant hybrid proteins. All fusion proteins should be able to enter the nucleus because of the native NLS present in the GAL4 binding domain (Silver et al. 1984) or the SV40 large T antigen nuclear localization signal that is fused to activation tagged proteins (Durfee et al. 1993). Other features contained in all vectors are the ColE1 origin of replication and the beta lactamase gene for replication and selection in *Escherichia coli*. For selection and replication in yeast, each vector contains the 2-micron origin and a nutritional marker. Table 2.1 lists the vectors used in this work, the markers they carry, and the vectors from which they were originally made.

name of vector	derived from	fusion	tag	eukaryotic selection	reference
pGAD-C(X)	pGAD424	GAL4 AD	/	LEU2	James et al. 1996
pGBD-C(X)	pGBT9	GAL4 DB	/	TRP1	James et al. 1996
pGBDU-C(X)	pGBT9	GAL4 DB	/	URA	James et al. 1996
pACT2		GAL4 AD	HA	LEU2	Durfee et al 1993
pODB8	pAS2	GAL4 DB	HA	TRP1	Louvet et al. 1997
pODB80	pAS2	GAL4 DB	/	TRP1	Louvet et al. 1997

Table 2.1 Description of vectors used in this project

pACT2 *Drosophila* embryonic expression library

The cDNA expression library used for the screen was made from mRNA derived from 0-18 hour embryos (gift from Steve Elledge). The cDNA pool was selected for transcripts greater than 1kb and cloned into the lamda phage vector λACT2 (activation domain, Durfee et al.1993) The λACT2 vector is converted from a phage to a plasmid using *cre-lox* mediated site-specific recombination. Thus, pACT2 is the plasmid excised from λACT2. Upon receipt, the *Drosophila* cDNA library was converted from phage into plasmid by subcloning conversion (see Materials and Methods for more details).

DB-Dof△C: The fusion protein used as bait

Instead of using the complete ORF of Dof as bait in the yeast two-hybrid screen, a C-terminal deletion of the protein was used. This is shown in Figure 2.2. This C-terminal deletion, is sufficient to rescue mesoderm and tracheal development in a *dof* mutant embryo (R. Wilson, personal communication). This deletion construct fused to the DNA binding domain is called DB-Dof△C and the identical region fused to the Gal4 activation domain is called AD-Dof△C.

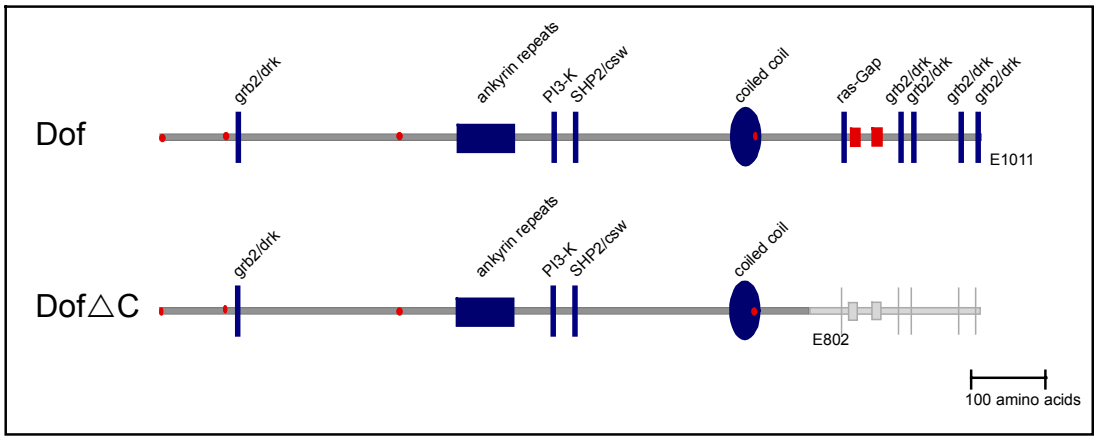


Figure 2.2 The upper figure is a drawing of the full length Dof protein (transcript II) adapted from Vincent et al. (1998). The ankyrin repeat is marked by a square and the coiled coil domain by an oval. Vertical bars indicate the position of tyrosines which, when phosphorylated, might be binding sites for the proteins indicated above the diagram. Red squares show two regions with putative PEST sequences (Rogers et al. 1986). Red dots are potential SUMOlation sites (Melchior 2000). The lower figure shows the Dof C-terminal truncation Dof△C. The deleted region is shown in light grey. E denotes the last amino acid and S denotes the first amino acid in the protein sequence

Tests with the DB-Dof△C construct

Before using a fusion protein as bait in a yeast two-hybrid screen, it is important to determine whether it is expressed, show that it enters the nucleus and check that it does not autoactivate, that is, activate transcription of reporter genes in the absence of the AD-hybrid protein. During initial test work, we observed that DB-Dof△C and AD-Dof△C interact in the yeast two-hybrid system. This provides a control for the behaviour of the bait construct, and shows that it is being expressed in *Saccharomyces cerevisiae* and entering the nucleus. When co-expressed with an empty vector, no reporter activation occurs, indicating that DB-Dof△C does not autoactivate. Table 2.1 shows the results.

	DB-Dof△C	pGBDU-c(1)
AD-Dof△C	+++	-
pGAD-c(1)	-	-

Table 2.2 The left column lists the activation domain fusion proteins and the top row contains the DNA binding domain fusion proteins. "+++" indicates growth on SC-His plates after 2 days at 30°C. "–" means that no growth occurred.

AR2.9: Dof△C integrated into PJ69-4a

For the yeast two-hybrid screen, the construct encoding DB-Dof△C was integrated into the chromosome of the host strain PJ69-4a. Two advantages have been observed when expressing a fusion protein from a chromosomally integrated construct rather than expressing a fusion protein from a plasmid. Integration of an expression construct eliminates plasmid recombination and allows a second construct (in this case, the expression library) to be transformed into the strain with a greater efficiency. The chromosomal integration of the Dof bait construct was accomplished by removing the 2-

micron sequence (origin of replication) from the DB-Dof△C plasmid and transforming pJG69-4a with a linearized plasmid. Upon selecting for an integrated DB-Dof△C , the only cells that replicate are those in which the cDNA has integrated itself into the chromosome, since the 2-micron is absent (see Materials and Methods). To ensure that the integration of the construct did not affect the expression of DB-Dof△C or of any other marker, Dof homodimerization was assayed in 10 of the strains in which integration had occurred. The strain with the highest transformation efficiency and showing the strongest activation of the reporter His3 with AD-Dof△C was chosen to be used in the yeast two-hybrid screen. This strain was called AR2.9.

Autoactivation test for fusion proteins

All fusion proteins that were made in this project were tested for autoactivation. For example, if an activation tagged protein was used, it was co-expressed with the "empty" expression vector, pGBDU. Hereafter, this control test will not be mentioned, unless the fusion protein autoactivated.

2.2 Description of the yeast two-hybrid screen

The yeast two-hybrid screen was carried out in two parts. Initially, His+lacZ+ activators were isolated and characterized. These candidates were sequenced and placed into groups of clones coding for the same gene. The longest transcript from each group was subsequently tested for activation of the Ade2 marker in the second part of the screen.

Selection for activation of His3 and lacZ markers

The *Drosophila* cDNA expression library (from S. Elledge, see above) was transformed into AR2.9 to yield a total of $1.3X10^6$ transformants. The transformants were grown on SC-ULH medium for 6-7 days, allowing for the growth of His3 positive clones. Next, the transformants were replica plated onto SC-ULH/Xgal plates and allowed to grow for 2-3 days. 269 colonies positive for His3 and lacZ were streak purified on SC-ULH plates and plasmid DNA was subsequently isolated. The plasmid DNA was checked by digesting the DNA with the restriction enzymes EcoRI and XhoI, which separates the insert from the vector. The plasmids were transformed back into AR2.9 to reconfirm the interaction. In parallel, the plasmids were transformed into PJ69-4a/pGBDU (a strain carrying an empty expression vector) and tested for autoactivation. Upon retransformation, 128 clones still activated the His3 reporter when co-expressed with DB-DofΔC but did not autoactivate. The 5 prime end of each of these clones was sequenced. The sequence data was used to search the databases BDGP, BLAST (see Materials and Methods) in order to identify transcripts that code for known genes as well as to identify transcripts that were isolated multiple times in the screen. The sequence data showed that the 128 clones consisted of 15 transcripts that were isolated multiple times, and 33 transcripts that were isolated only once. The list of these clones is shown in Table 2.3. 2 single clones were subsequently excluded from further analysis because one sequence (i99) contained a chain of 26 "GAG" triplets and the other clone (i30) was derived from mitochondrial 16SRNA. Thus the final count of his+lacZ+ clones comes to 15 groups and 31 single copies.

Table 2.3 This table is a list of the transcripts isolated in the yeast two-hybrid screen. Clones that activate the Ade2 reporter are highlighted in yellow. Information about motifs and homologies is given for genes that have not been described by a name in Flybase.(1) The name of the gene is given to clones whose 5 prime sequence is identical to the transcript of a gene that has been given a name by FlyBase (see Materials and Methods) in a FlyBase Report. Note that this can also be a putative gene that has not been characterized in *Drosophila* but whose sequence annotation is homologous to a gene that has been characterized in another species. (2) CG14206: residues 1–127 share 62% amino acid identity with residues 3–130 of Human 40S ribosomal protein S10 (3) The motif found using the PFAM database (see Materials and Methods) is noted here if the transcript has not been given a name by Flybase (4) If the clone is identical to an annotated gene from the Genome Annotation Database at Flybase (see Materials and Methods), the CG number is listed. If the clone is not identical to any annotated gene, the genomic sequence to which it maps (Adams M.D. et al 2000) is denoted with an AE number**.**

motifs (3) or aa homology (2)	gene (1)	members	longest clone	CG#/ AE# (4)
fn3 (3)		21	i119	CG 12340
	Acon	13	i59	CG 9244
Human 40S ribosomal protein S10(2)		10	i14	CG 14206
	dof	9	i28	CG 3375
		6	i238	CG 2713
	ubc9	5	i56	CG 3018
	bap60	5	i133	CG 4303
	heartless	5	i150	CG 7223
cyclin domain (3)		4	i220	CG 11525
	βspectrin	4	i6	CG 5870
DH domain, PH domain (3)		3	i19	CG 2008
		3	i159	AE 003492
		2	i199	CG 10189
		2	i7	AE003433
	bip2	2	i163	CG 2009
	rack1	1	i173	CG 7111
	sinuous	1	i8	CG 10624
	GAP69C	1	i51	CG 4237
	cup	1	i239	CG 11181
	PP2A –B'	1	i234	CG 7913
	msl-1	1	i94	CG 10385
	hr39	1	i70	CG 8676
	Dmzimp	1	i184	CG 8068
	haywire	1	i50	CG 8019
	RpA-70	1	i164	CG 9633
	not	1	i155	CG 4166
	BC	1	i12	CG 5779
	Eps - 15	1	i31	CG 16932
ankyrin repeats, zinc finger (3)		1	i115	CG 17492
WD40 domains (3)		1	i207	CG 14805
	acGAP	1	i249	CG 13345
zinc finger, C2H2 type (3)		1	i36	CG 11762
BTB/POZ domain (3)		1	i188	CG 11275
ABC transporter (3)		1	i17	CG 12703
	Hsc70-4	1	i190	CG 4264
SPRY domain (3)		1	i25	CG2944
Acetyl-CoA hydrolase (3)		1	i130	CG 7920
		1	i118	AE 003512
		1	i136	CG 10671
		1	i175	CG 15740
		1	i182	CG 14939
		1	i232	AE 003809
		1	i169	AE 003512
		1	i176	AE 003633
		1	i226	CG 4147
		1	i187	no hits
		1	i99	
		1	i30	

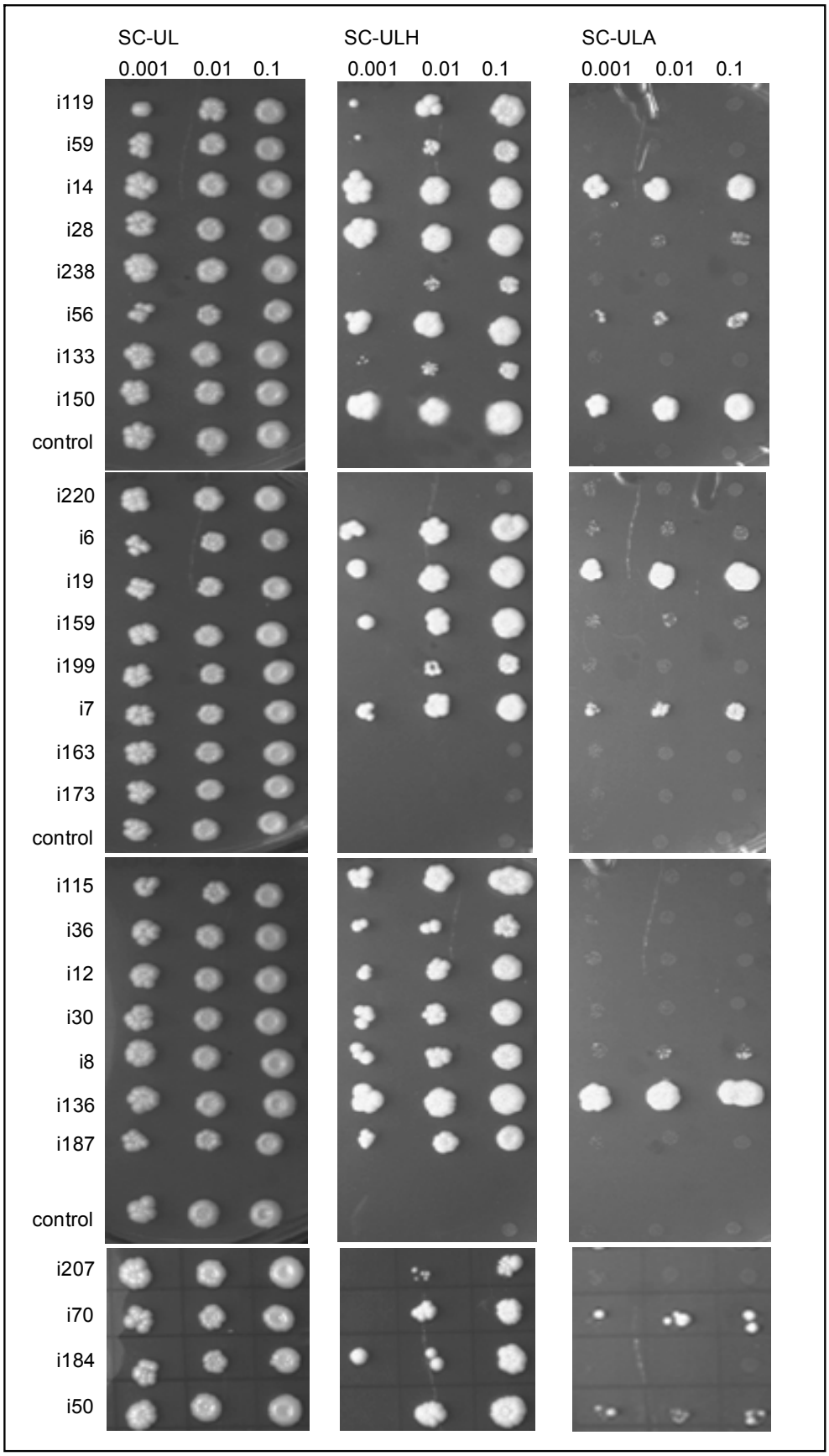
Selection of Ade+ interactions

In the final round, all candidates that activated the His3 reporter were co-expressed with DB-DofΔC to test whether the interaction was strong enough to activate the Ade2 reporter, which has the most stringent promoter of all three reporter genes (see above). Eight groups and 4 single clones activated Ade2. The clones that activated the Ade2 reporter are highlighted in Table 2.3 and the co-expressions are shown in Figure 2.3.

Binding of candidates to full length Dof and to a Dof deletion construct

All candidates were tested for activation of the His3 reporter when co-expressed with the full length Dof construct and the deletion construct 430, which does not rescue Dof function in *dof* mutant embryos (R. Wilson, personal communication). These results are shown in Appendix II.

Figure 2.3 In this figure, candidates from the yeast two-hybrid screen were co-expressed with DB-DofΔC. The control in all cases is the empty vector pGBDU co-expressed with i150. The only difference between this test and the test documented in Table 2.3 is that here, a plasmid-borne DB-DofΔC is used in the co-expression experiment whereas the tests for activation of the His3 reporter (summarized in Table 2.3) were carried out with the construct DB-DofΔC that was integrated into chromosomal DNA. This might be the reason for the fact that the clones i173 and i163 don't activate His3 when co-expressed with the plasmid-borne DB-DofΔC. Plates on which the interacters grew on SC-ULH but not on SC-ULA are not shown. Method used for "dot assay": A single colony was inoculated in selection liquid medium (in which the presence of the plasmid is selected) and grown at 30°C for 24 hours. 1µl from a dilution series was pipetted onto selection plates as indicated. The pictures were taken after plates were incubated for 5.5 days at 30°C.


SC-UL
SC-ULH
SC-ULA
0.001 0.01 0.1
0.001 0.01 0.1
0.001 0.01 0.1
i119
i59
i14
i28
i238
i56
i133
i150
control
i220
i6
i19
i159
i199
i7
i163
i173
control
i115
i36
i12
i30
i8
i136
i187
control
i207
i70
i184
i50

2.3 Heartless

Five transcripts corresponding to the intracellular domain of Heartless were isolated in the screen. The sequence data showed that all transcripts are in-frame fusions with the Gal4 activation domain. Figure 2.4A shows the positions of the first amino acid of each clone in relation to the amino acid sequences of Heartless, Breathless and FGFR1. The longest transcript, i150, contains the whole intracellular domain of Heartless. The shortest transcript, i79, contains only two thirds of the protein kinase domain and shows that only the C-terminal part of the intracellular domain of the Heartless receptor is essential for binding to DofΔC. The crystal structure of the cytoplasmic domain of human FGFR1 (Mohammadi et al. 1996) is depicted by a ribbon diagram in Figure 2.4B and 2.4C. Assuming that the structures of Heartless, Breathless and FGFR1 are conserved, the structure of FGFR1 can be used to illustrate the regions of the Heartless that are contacting Dof. The longest clone, i150, contains both lobes of the intracellular domain and the shortest clone, i79, only contains the bottom lobe.

Breathless binds to Dof in the yeast two-hybrid system

As mentioned in the Introduction, it has been genetically shown that Dof lies downstream of both known Drosophila FGF receptors, namely Heartless and Breathless. However, only transcripts of *Heartless* were isolated in this screen. In order to test whether Dof also interacts with Breathless in the yeast two-hybrid system, the cDNA corresponding to minimal Heartless sequence that binds to Dof (that is, construct i79) was cloned into yeast two-hybrid vectors. The construct is called Btl-K

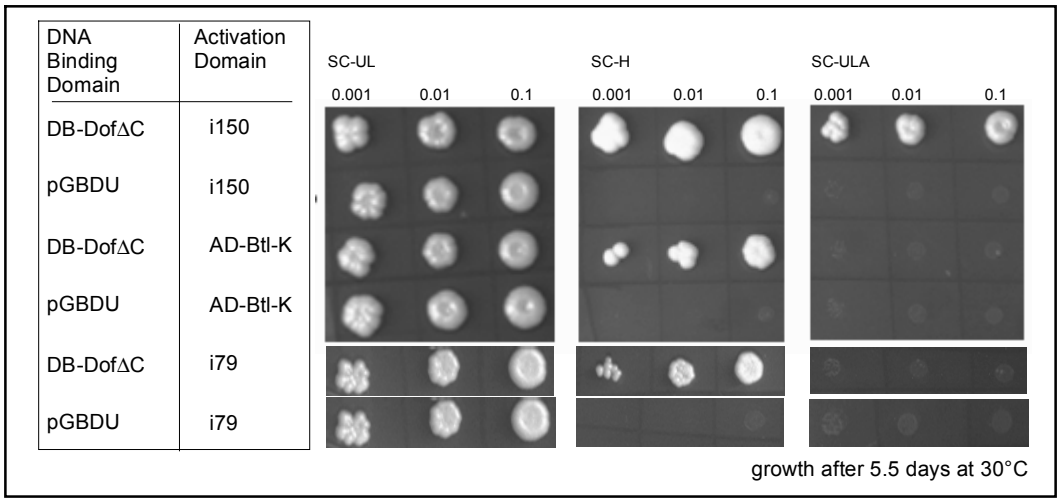


Figure 2.5 The activation domain fusion with Breathless (AD-Btl-K) activates the reporter His3 when co-expressed with DB-DofΔC. However, like the equivalent sequence of Heartless (i79), AD-Btl-K does not activate the Ade2 reporter when co-expressed with DofΔC. In contrast, the longest sequence of Heartless isolated in the screen (i150) does activate Ade2 when co-expressed with DofΔC. See Figure 2.3 for description of dot assay.

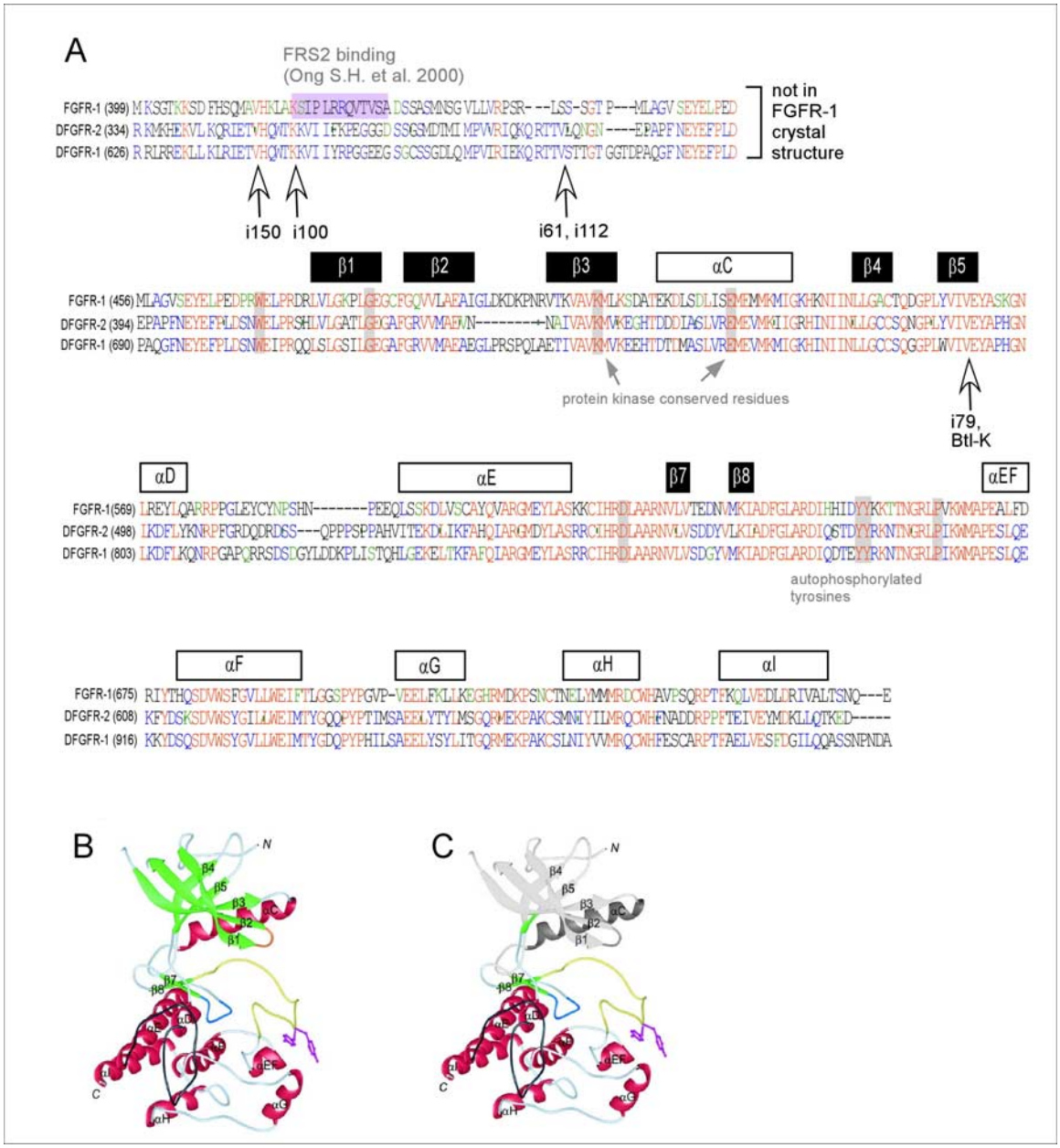


Figure 2.4 **A** is an alignment of cytoplasmic domains of FGFR-1 with Heartless and Breathless and arrows point to the first amino acid of each transcript isolated in the yeast two-hybrid screen. The Helix and Beta-sheet positions are for FGFR-1 and taken from Mohammadi et al. 1996. DFGFR-1: Breathless;DFGFR-2: Heartless; The amino acid sequence of DFGR-2 is deduced from CT39172. **B** and **C** are ribbon diagrams of the cytoplasmic FGFR1 structure, which does not include the juxtamembrane region and the last few residues of the C-terminus (adapted from Mohammadi et al. 1996). The a helices are shown in red, the b strands in green, the nucleotide-binding loop in orange, the catalytic loop in blue, the activation loop in yellow, the kinase insert in black, and the side chains of Tyr-653 and Tyr-654 in purple. In B, the region that is colored is equivalent to that encoded by i79, the shortest Heartless transcript isolated in the yeast two-hybrid screen.

As Figure 2.4A shows, the construct Btl-K contains most of the protein kinase domain but lacks the ATP binding region. The co-expression of AD-Btl-K with DB-Dof△C

resulted in an activation of His3, but not of Ade2 (Figure 2.5). However, i79, the equivalent Heartless construct does not activate the Ade2 either. This result shows that although only Heartless was isolated in the yeast two-hybrid screen, Dof△C binds to Breathless as well in the yeast two-hybrid system.

The cytoplasmic domain of Heartless only interacts with Dof in one direction

Only the activation-tagged cytoplasmic domain of Heartless interacts with DB-Dof△C (Figure 2.6). The Gal4 DNA-binding domain fusion with cytoplasmic Heartless (DB-Htl) does not activate His3. One reason for this may be that several intracellular receptor domains are brought together at the Gal4 DNA binding sites, resulting in their autophosphorylation. It is therefore conceivable that Dof does not bind to the phosphorylated receptor and only binds to the unphosphorylated receptor, because when the activation domain is fused to the cytoplasmic domain of Heartless, this domain can bind to Dof. This idea is supported by the fact that the Breathless construct DB-Btl-K weakly activates His3 (Figure 2.6). DB-Btl-K only contains two thirds of the protein kinase domain and does not contain the ATP-binding site. Thus, this construct doesn't have a functional kinase domain and when the DB-Btl-K fusion proteins are brought together at the Gal4DNA binding sites, they cannot phosphorylate each other.

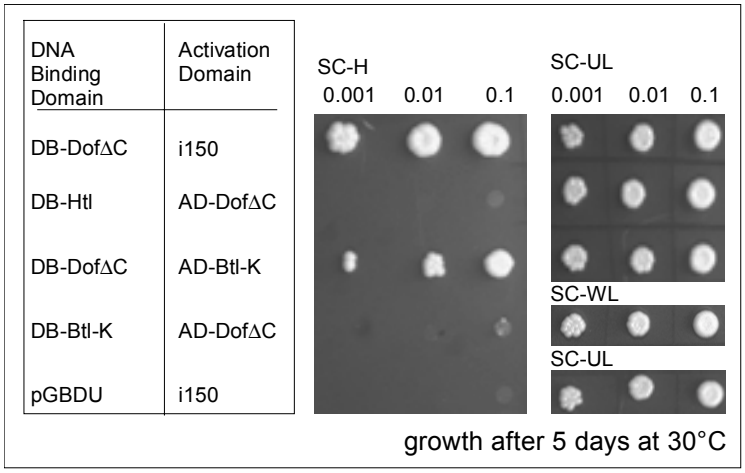


Figure 2.6 The cytoplasmic domain of Heartless only interacts with DofΔC when it is fused to the activation domain (i150). When it is fused to the DNA-binding domain (DB-Htl) it does not activate His3 when co-expressed with DofΔC. When a short version of cytoplasmic Breathless, which lacks the ATP binding site, is fused to the DNA binding domain and co-expressed with DofΔC, slight activation is observed. See Figure 2.3 for description of dot assay.

i79 and AD-Btl-K behave similarly with the Dof deletion panel

The Breathless construct AD-Btl-K was tested against a panel of Dof deletions to determine whether it behaved similarly to i79, the shortest Heartless construct. Similarly, the full length cytoplasmic domain of Heartless was tested against the Dof deletion panel. These tests are shown in Figure 2.7.
AD-Btl-K and i79 have the same binding behaviour with Dof. Interestingly, unlike construct i150 (the full cytoplasmic domain of Heartless), i79 and Btl-K do not interact as

strongly with full length Dof as they do with some of the deletions. Furthermore, DB430, containing a deletion in Dof which abolishes its binding to i150, still shows weak interaction with i79 and Btl-K.

Full length Dof binds to the full length cytoplasmic domain of Heartless (i150) but its C-terminus appears to diminish the strength in binding because when the co-expression with i150 is tested for the activation of Ade2 (the more stringent reporter), the full length Dof construct does not activate Ade2 transcription, but DofΔC does. More than half the Dof molecule can be deleted from the C-terminus without affecting its binding to Heartless (construct 405). However, sequences in the N-terminal half are required for binding to Heartless because large deletions (DB354, DB264) as well as internal deletions (DB430) affect binding. DB427 represents the largest N-terminal deletion that will still interact with Dof. However, this statement is only partly valid, since DB427 is deleted at the C-terminus as well, and was used here because the construct with the C-terminus auto-activates.

A single amino acid change abolishes the interaction between Dof and the FGFRs

As can be seen in Figure 2.7, deletions of portions of the Dof N-terminus diminish its ability to bind to the full cytoplasmic domain of Heartless (i150). In fact, a single amino acid exchange, namely from a Lysine to an Arginine at residue 296 in Dof, completely abolishes binding of Dof to all the receptor constructs (i150, i79 and Btl-K) when assayed for activation of the His3 reporter. This is shown in figure 2.17 in Chapter 2.5. The lysine that was mutated is a potential SUMO acceptor site, indicating that either SUMOlation of Dof is required for interaction with the full cytoplasmic domain, or that the Lysine to Arginine mutation resulted in a structural change of the receptor binding site. In the least, this Lysine has been shown to be essential for the interaction of Dof with the FGFRs in the yeast two-hybrid system.

Oddly, the minimal receptor constructs (i79 and AD-Btl-K) do still activate the His3 reporter when co-expressed with Dof molecules that lack the potential SUMO acceptor lysine (constructs 430 and 264). This contradicts the observation that a single amino acid exchange in this region abolishes binding of Dof with the receptors. This might point to structural changes that have occurred due to the lack of a SUMOlation site or the substitution of Lysine for Arginine.

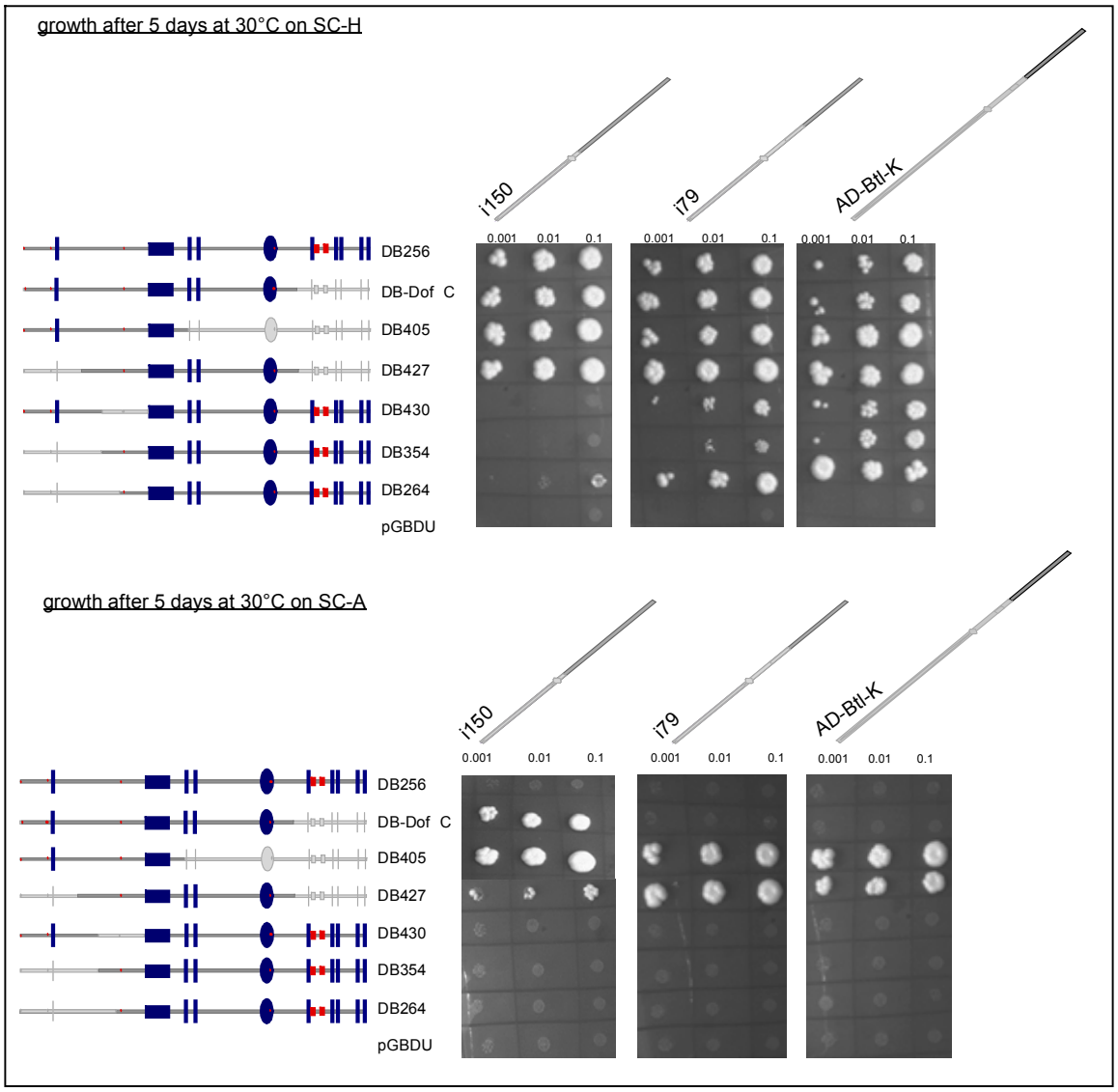


Figure 2.7 This figure shows the interaction of the FGFR constructs (i150, i79 and AD-Btl-K) with a panel of Dof deletion constructs. The interaction was assayed with the reporters His3 (top row of pictures) and Ade2 (bottom row of pictures). The Dof deletion constructs are drawn here with deleted regions shaded in grey (see Materials and Methods for more details and Figure 2.1 for a description of motifs). The drawings of the FGFR constructs also show which part of the receptor sequence is included in the construct. The circle shows the position of the transmembrane domain of the receptors. See Figure 2.3 for description of dot assay. As a control, all co-expressions were simultaneously tested on SC-UL selection plates. All co-expressions grew, but the data is not shown here.

2.4 Dof

Nine Dof transcripts were isolated in the yeast two-hybrid screen. Sequence analysis of the 5 prime end showed that all transcripts are in-frame fusions with the Gal4 activation domain. The protein sequences of the longest and shortest transcript are aligned with that of full length Dof in Figure 2.8

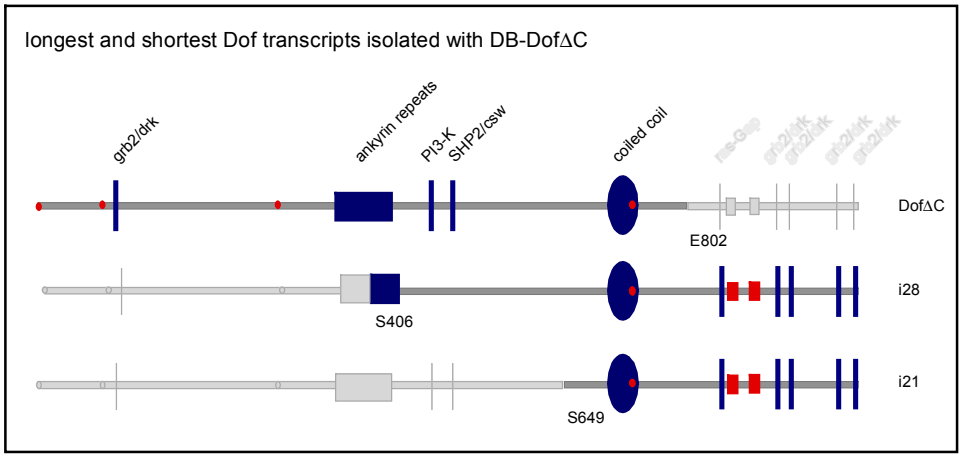


Figure 2.8 This figure shows DofΔC aligned with the shortest (i21) and longest (i28) Dof transcript isolated in the yeast two-hybrid screen. Areas shaded in light grey indicate regions that are absent from the protein. E denotes the last residue and S denotes the first residue in the protein sequence. See Figure 2.2 for a description of motifs.

None of the nine transcripts isolated in the yeast two-hybrid screen code for the full length Dof protein, each lacking sequence from the N-terminus. Figure 2.9A illustrates the interaction between DofΔC and the shortest Dof transcript isolated in the screen.

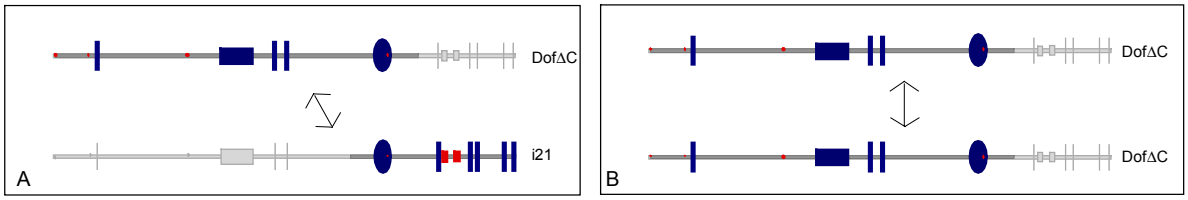


Figure 2.9 A. Interaction data from the yeast two-hybrid screen show that an N-terminal deletion of Dof can interact with a C-terminal deletion of Dof. **B**. Two C-terminal deletion constructs of Dof can interact in the yeast two-hybrid system

In a separate set of experiments, co-expression of AD-DofΔC and DB-DofΔC in the yeast two-hybrid system resulted in the activation of all three reporters, His3, lacZ and Ade2, and indicated that the C-terminal deletion of Dof also binds to itself in the yeast two-hybrid system. This is illustrated in Figure 2.9B. These yeast two-hybrid interactions provide potential clues as to how Dof might be functioning in its natural environment. One hypothesis is that Dof functions as a homodimer. A second hypothesis (which does not exclude the former hypothesis) is that intramolecular interactions between distinct

domains of Dof regulate its behaviour. For example, the affinity of Dof to other binding partners might be modulated by changing conformations via intramolecular interactions.

Mapping the dimerization domains in Dof

In order to determine which regions of Dof interact, a panel of Dof deletion constructs was cloned into the yeast two-hybrid expression vectors pGAD and pGBDU/pGBD, such that each deletion construct could be expressed as an activation domain fusion protein or as a binding domain fusion protein. As an example of a co-expression, Figure 2.10 depicts full length Dof fused to the Gal4 DNA-binding domain and full length Dof fused to the Gal4 activation domain.

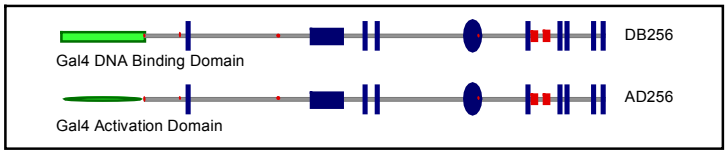


Figure 2.10 This figure is an example two constructs used in a co-expression experiment. Here, one full length Dof transcript is fused to the Gal4 DNA-binding domain sequence and another full length Dof transcript is fused to the Gal4 activation domain. In the pGBDU/pGBD-pGAD vectors used for these assays, all constructs will have the fusion moiety at the N-terminus.

Method used for scoring an interaction between two deletion constructs

Pairs of deletion constructs were co-expressed in PJ69-4A and tested for activation of the His3 reporter after growth on SC-His agar plates for four and a half days at 30°C. Since different rates of growth were observed, the activation of the His3 reporter was scored in the following manner: If co-expression led to colonies as large as those observed with DofΔC dimerization (AD-DofΔC + BD-DofΔC), then the interaction was scored as a "+". If co-expression led to colonies smaller than DofΔC dimerization but larger than the negative control (DofΔC + empty AD vector) it was scored with a "+". If there was no difference between the co-expression culture and the negative control, then this co-expression was scored as a "0". Figure 2.11 shows a picture of a typical co-expression experiment and the scores that each co-expression would have received.

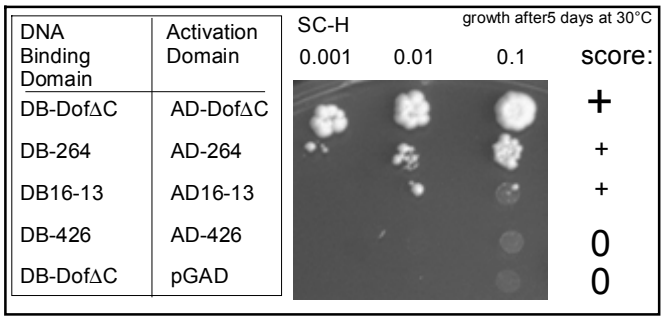


Figure 2.11 How co-expressions were scored in this experiment. See text for details

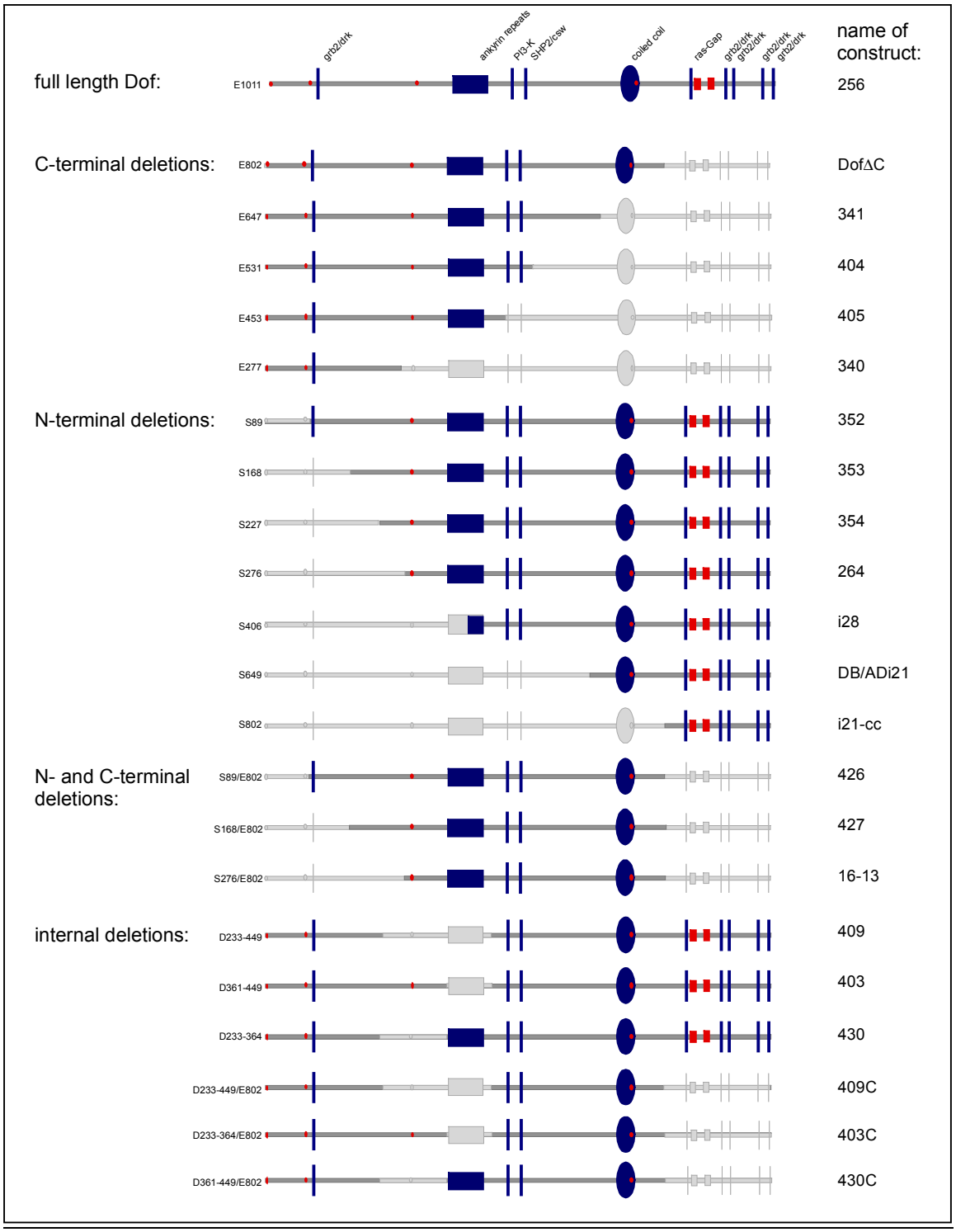


Figure 2.12 This figure shows all the constructs that were used to establish regions of Dof that are homodimerizing. Areas shaded in light grey indicate regions that are absent from the fusion protein. E denotes the last amino acid in a sequence, S denotes the first amino acid in a sequence, and D denotes the first and last amino acid deleted from an interior region. Note that all constructs were cloned into the pGBDU/pGBD-pGAD vectors, except for i28, which is in the vector pACT. See materials and methods for cloning details and Figure 2.2 for a description of motifs.

Figure 2.12 lists the Dof deletion constructs used for the domain mapping experiment. The figure organizes the deletion constructs into five categories: Full length, N-terminal deletions, C-terminal deletions, N-and C- terminal deletions, and internal deletions. The results of co-expressing these constructs are shown in Table 2.3 and the results will be discussed below.

	256	265	341	404	405	340	352	353	354	264	16/13	426	427	409C	409	403C	403	430C	430	DBi21	i21cc	pGBD
256	+	+		+	+	0	+	+	+			+	+	0	+	+		+	0	+	0	0
265	+	+	+	+	+	0	+	+	+	0	0	0	+		+	+		0		+	0	0
341	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
404	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
405	+	+	+	+	+	0	+	+	+	+		0	+	0	+	+	+	0	+	+		0
340	0	0			0	0							0	0						0		0
352	+	+			+		+													0	0	0
353	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
354	+	+			+				+											+	0	0
264		+			+					+										+	+	0
16/13		0									+											0
426	+	+			0							0								0	0	0
427	+				+	0							+	0					+	+	0	0
409C	0				0	0							0	0	+	+		0		0	0	0
409	+	+			+									+	+	+		+		0		0
403C	+	+			+									0	0	+						0
403	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
430C	0	0			0									0	0			0		0		0
430	+				+			+					+						0	0		0
DBi21	0	+	+	+	+	0	0	0	0	0	+	+	0	0	0	+	0	0	0	0	0	0
i21cc	0	+			0		0	0	0		0	0	0	0						0	0	0
pGBD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0

Table 2.3 This table shows the scores given to the activation of His3 upon co-expression in PJ694a of two Dof fusion proteins containing the series of deletions depicted in Figure 2.12. The first column (shaded blue) contains the name of the construct cloned into the DNA-binding domain vector and the top row (shaded yellow) lists all constructs cloned into the activation domain vector. See materials and methods for cloning details. The green boxes marks the co-expression of an identical construct. If the construct autoactivates, the co-expression is marked with a "/". All other scores are described in Figure 2.11.

Interpretation of the Co-expression of Dof deletion constructs

Before the interpretation is presented, two limitations of the yeast two-hybrid system in characterizing an interaction should be pointed out. First if two proteins interact in one orientation, they might not interact when they are in the other orientation. For example if a protein interacts with Dof as a DNA binding domain fusion, it might not interact with Dof as an activation domain fusion. This phenomenon was observed several times in the co-expression experiments done here (see for example construct 403C and construct 409 co-expression). Second, identical proteins expressed from different vectors will exhibit different binding behaviour and the construct i21 which was isolated in the yeast two-hybrid screen, does not interact very strongly with DofΔC when expressed from the pGBDU/pGAD vector series, which was used to express all the deletion constructs (see Appendix I). Thus it is important to keep in mind that besides structural changes might be caused by deleting large regions of the Dof molecule, yeast-two hybrid interactions can only be interpreted to a certain point.

Any Dof molecule containing DR#1 homodimerizes

Construct 405 codes for the N-terminal third of Dof. This construct self-associates and defines an N-terminal region that can homodimerize. For this discussion, the region spanning amino acids 1-453 will be defined as "Dimerization Region #1 (DR#1). The interaction datum predicts that any Dof molecule containing DR#1 will homodimerize. This is illustrated in figure 2.12A. Table 2.4 highlights all co-expressions that confirm or contradict this hypothesis.

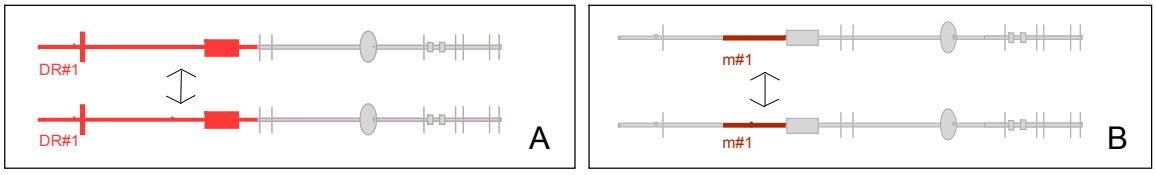


Figure 2.12

The next step is to narrow down the dimerization region of DR#1 by examining the behaviour of constructs containing deletions in DR#1. The construct 430C is as long as DofΔC, but is missing amino acids 233-634. Whereas DofΔC dimerizes, construct 430C does not. Construct 430 allows to define a module "m#1" (residues 361-441) within DR#1 that is required for DR#1 dimerization (Figure 2.12B). It remains to be seen whether module m#1 can dimerize on its own. This would prove that it is this region alone that is mediating the dimerization of DR#1. Table 2.4 highlights supporting and contradicting results. Note that construct 340 is even shorter than DR#1 and does not homodimerize but will be excluded from this analysis since there is no evidence that this construct is being expressed (it does not interact with any other construct).

			m1	m1	m1	m1	m1	0.5 m1	m1	m1	m1	0.5 m1	0.5 m1	m1	m1			m1	m1					
		DR #	256	265	341	404	405	340	352	353	354	264	16/13	426	427	409C	409	403C	403	430C	430	i21	i21-cc	pGAD
	DR #		1,2	1	1	1	1		2	2	2	2					2		2		2	2		
m1	256	1,2	+	+		+	+	0	+	+	+			+	+	0	+	+		+	0	+	0	0
m1	265	1	+	+	+	+	+	0	+	+	+	0	0	0	+		+	+		0		+	0	0
m1	341	1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	404	1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	405	1	+	+	+	+	+	0	+	+	+	+		0	+	0	+	+	+	0	+	+		0
0.5 m1	340		0	0			0	0							0	0						0		0
m1	352	2	+	+			+		+													0	0	0
m1	353	2	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	354	2	+	+			+				+											+	0	0
0.5 m1	264	2		+			+					+										+	+	0
0.5 m1	16/13			0									+											0
m1	426		+	+			0							0								0	0	0
m1	427		+				+	0							+	0					+	+	0	0
	409C		0				0	0							0	0	+	+		0		0	0	0
	409	2	+	+			+									+	+	+		+		0		0
m1	403C		+	+			+									0	0	+						0
m1	403	2	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
	430C		0	0			0									0	0			0		0		0
	430	2	+				+			+					+						0	0		0
	i21	2	0	+	+	+	+	0	0	0	0	0	+	+	0	0	0	+	0	0	0	0	0	0
	i21cc		0	+			0		0	0	0		0	0	0	0						0	0	0
	pGBD		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0

Table 2.4 This table contains the same data as Table 2.3, but highlights the co-expressions that support the hypothesis that any Dof molecule containing DR#1 will homodimerize. Co-expressions in support of this hypothesis are coloured in red: "+". Dof deletion constructs that only need module m#1 in DR#1 to homodimerize are coloured in turquoise: " + ". Co-expressions shaded in grey are those that contradict the model in one direction. Co-expressions shaded in grey stripes are those that contradict the model in both directions. Additional columns / rows are included here to point out which constructs contain the domains m#1 (called "m1" here), DR#1 and DR#2 (called "1" and "2" here).

Any Dof molecule containing DR#2 will bind to any Dof molecule containing DR#1

Construct AD/DBi21 codes for the C-terminal third of the Dof molecule and binds to construct 405. Construct DB/ADi21 codes for residues 649-1011 and defines the "Dimerization Region#2" (DR#2). DB/ADi21 does not bind to itself, but binds to any construct that contains DR#1. On the basis of these results, two further statements can be made: Any molecule containing DR#2 will bind to any Dof molecule containing DR#1. DR#2 on its own, however, is not sufficient for dimerization (Figure2.13). Results confirming or negating these statements are highlighted in Table 2.5.

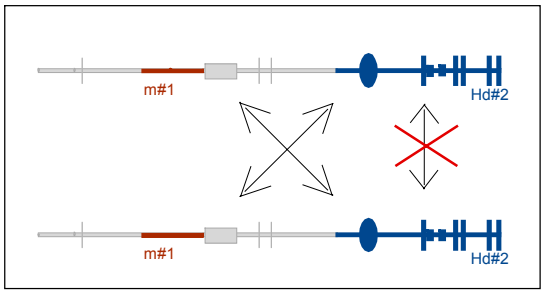


Figure 2.13

Is m#1 the same module within DR#1 that is binding to DR#2? This is very likely since 403C does not bind to i21, whereas the neighboring deletion, 430C, does bind to DB/AD i21. Put together, the model predicts that the Dof molecule will self-associate because m#1 is required for DR#1 to bind to itself and to DR#2. The model is shown in Figure 2.14.

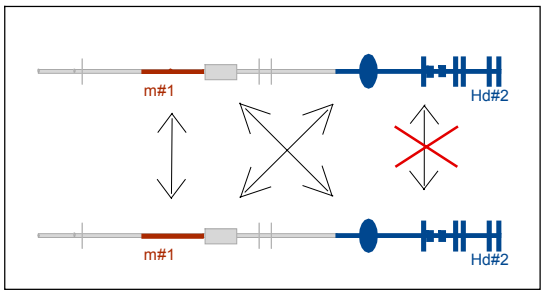


Figure 2.14

The putative SUMOlation site in m#1 is not required for Dof self-association

Since it has been observed that SUMO moieties homodimerize in yeast (see Chapter 2.5), and since the third putative SUMOlation site is located in the region m#1, the hypothesis that Dof might dimerizes because it is SUMOlated, was tested. However, when this site was mutated, the molecules still dimerized (Figure 2.16, Chapter 2.5), thus excluding the possibility that homodimerization is mediated by a SUMOlation of Dof at its third putative SUMOlation site.

			m1	m1	m1	m1	m1	0.5 m1	m1	m1	m1	0.5 m1	0.5 m1	m1	m1			m1	m1					
		DR#	256	265	341	404	405	340	352	353	354	264	16/13	426	427	409C	409	403C	403	430C	430	i21	i21-cc	pGAD
	DR#		1,2	1	1	1	1		2	2	2	2					2		2		2	2		
m1	256	1,2	+	+		+	+	0	+	+	+			+	+	0	+	+		+	0	+	0	0
m1	265	1	+	+	+	+	+	0	+	+	+	0	0	0	+		+	+		0		+	0	0
m1	341	1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	404	1	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	405	1	+	+	+	+	+	0	+	+	+	+		0	+	0	+	+	+	0	+	+		0
0.5 m1	340		0	0			0	0							0	0						0		0
m1	352	2	+	+			+		+													0	0	0
m1	353	2	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
m1	354	2	+	+			+				+											+	0	0
0.5 m1	264	2		+			+					+										+	+	0
0.5 m1	16/13			0									+											0
m1	426		+	+			0							0								0	0	0
m1	427		+				+	0							+	0					+	+	0	0
	409C		0				0	0							0	0	+	+		0		0	0	0
	409	2	+	+			+									+	+	+		+		0		0
m1	403C		+	+			+									0	0	+						0
m1	403	2	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	+
	430C		0	0			0									0	0			0		0		0
	430	2	+				+			+					+						0	0		0
	i21	2	0	+	+	+	+	0	0	0	0	0	+	+	0	0	0	+	0	0	0	0	0	0
	i21cc		0	+			0		0	0	0		0	0	0	0						0	0	0
	pGBD		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0

Table 2.5 Coexpression results supporting the hypothesis that any Dof deletion construct containing DR#2 will bind to any Dof deletion construct containing DR#1 are coloured in red: "+" The co-expressions shaded in dark grey contradict the hypothesis. Co-expressions supporting the hypothesis that any Dof deletion construct containing DR#2 will bind to any Dof deletion construct containing m#1are coloured in brown: " + ". Co-expressions shaded in light grey don't fit the aforementioned hypothesis. The co-expressions that show that DR#2 is not sufficient for dimerization are coloured in pink: " 0 ". Additional columns / rows are included here to point out which constructs contain the domains m#1 (called "m1" here), DR#1 and DR#2 (called "1" and "2" here).

2.5 DmUbc9

Five transcripts from *Drosophila ubiquitin-conjugating enzyme 9* (*DmUbc9*) were isolated in the screen. Sequence analysis of the 5 prime end of each clone showed that all are in-frame fusions with the Gal4 activation domain. The longest transcripts, i77 and i105 are one amino acid shorter than the full length DmUbc9 and the shortest transcripts are 5 amino acids shorter than the full length DmUbc9 protein sequence. *DmUbc9* is synonymous with *lesswright*, abbreviated as *lwr*. *DmUbc9* encodes a SUMO-conjugating Enzyme. SUMO (small ubiquitin-related modifier) is a member of the family of ubiquitin-related proteins and, in contrast to ubiquitin, has been implicated in the regulation of protein-protein interactions and subcellular localization (reviewed in Melchior, 2000).

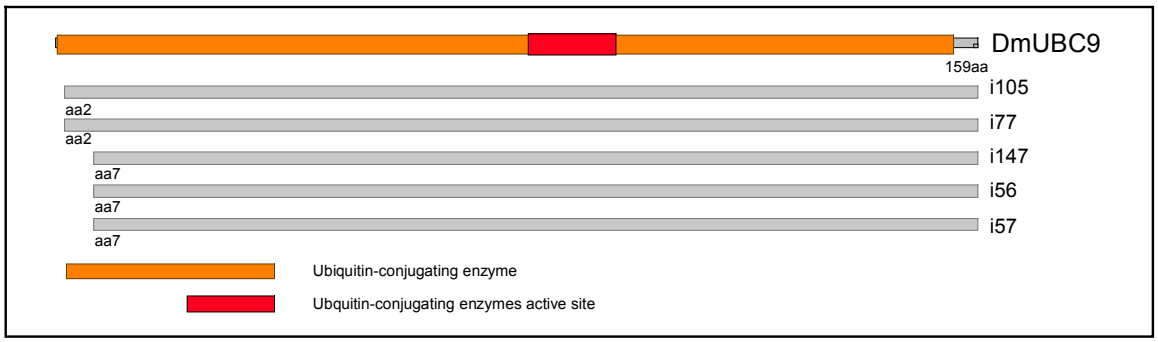


Figure 2.15 This figure shows the full length protein encoded by *DmUbc9* (from the accession number AB017607) aligned with the predicted amino acid sequence of the transcripts isolated in the yeast two-hybrid screen. The motifs shown in DmUBC9 are names of PFAM motifs (see Materials and Methods) scoring higher than gathering threshold. Note that the longest clones are i105 or i77. However, in all assays, the clone i56 was used to represent this group and thus misnamed "the longest clone".

Dof as a potential SUMO target

The minimal consensus sequence for SUMOlation is aKX(E,D) (Melchior, 2000) and Dof has four Lysines that could potentially serve as an acceptor site for the SUMO protein (Figure 2.16).

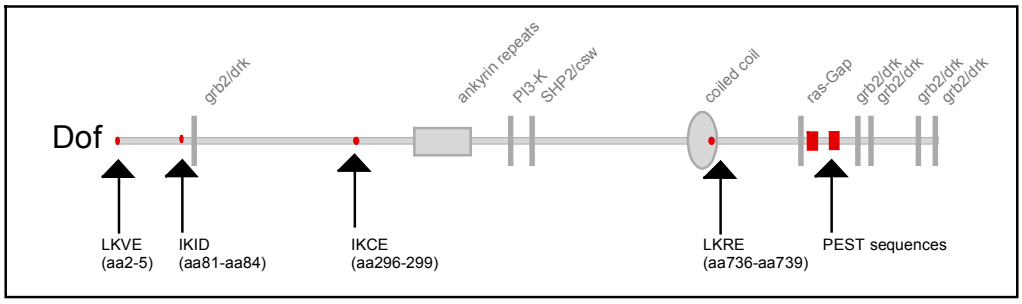


Figure 2.16 This is a picture of full length Dof (see Figure 2.2 for details and motifs) which points out the four putative SUMOlation sites and the PEST sequence.

Another feature that Dof shares with SUMO targets is its PEST sequence. When 19 verified SUMO targets were examined (Melchior 2000), 10 contained one or more PEST sequences.

Mutation of the third putative SUMOlation site in Dof

The third putative SUMOlation site in Dof (at residues 296 – 299) is located in a region of the molecule that when missing, affects Dof's *in vitro* and *in vivo* behaviour. When this region is deleted (D233-449), as for example in construct 430C, dimerization no longer occurs (Table 2.3). When D233-449 is deleted from the full length molecule (construct 430), interaction with some of the yeast two-hybrid candidates is compromised or does not occur at all anymore (Appendix II). Above all, construct 430 does not rescue the Dof mutant phenotype in the mesoderm and the trachea (R.Wilson, personal communication). Since yeast SUMO self-associates (I. Schwienhorst, personal communication), it is possible that if the third putative SUMOlation site in Dof is indeed SUMOlated, Dof dimerizes because of the SUMO moieties attached to it. In order to test this hypothesis, the third SUMOlation site in Dof was mutated by exchanging the Lysine (K297) for an Arginine. When a protein becomes SUMOlated, the COOH-group of the C-terminal Glycine of SUMO becomes covalently bound to the epsilon NH3-Group of the Lysine in the SUMOlation motif (reviewed in Melchior 2000). The SUMOlation site mutation was introduced into the DofΔC construct in order to assay dimerization of the N-terminal region Hd#1 without the disturbance of the C-terminus (region Hd#2) to which Hd#1 can also bind (see Chapter 2.4 for details). The construct was called "Δs.site3".

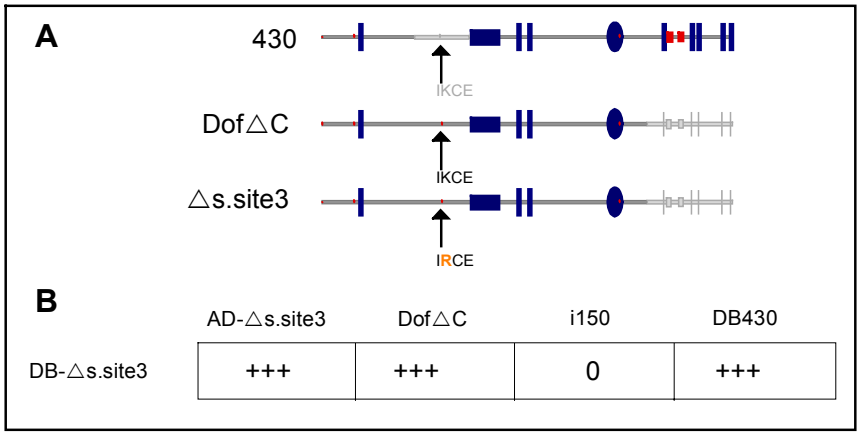
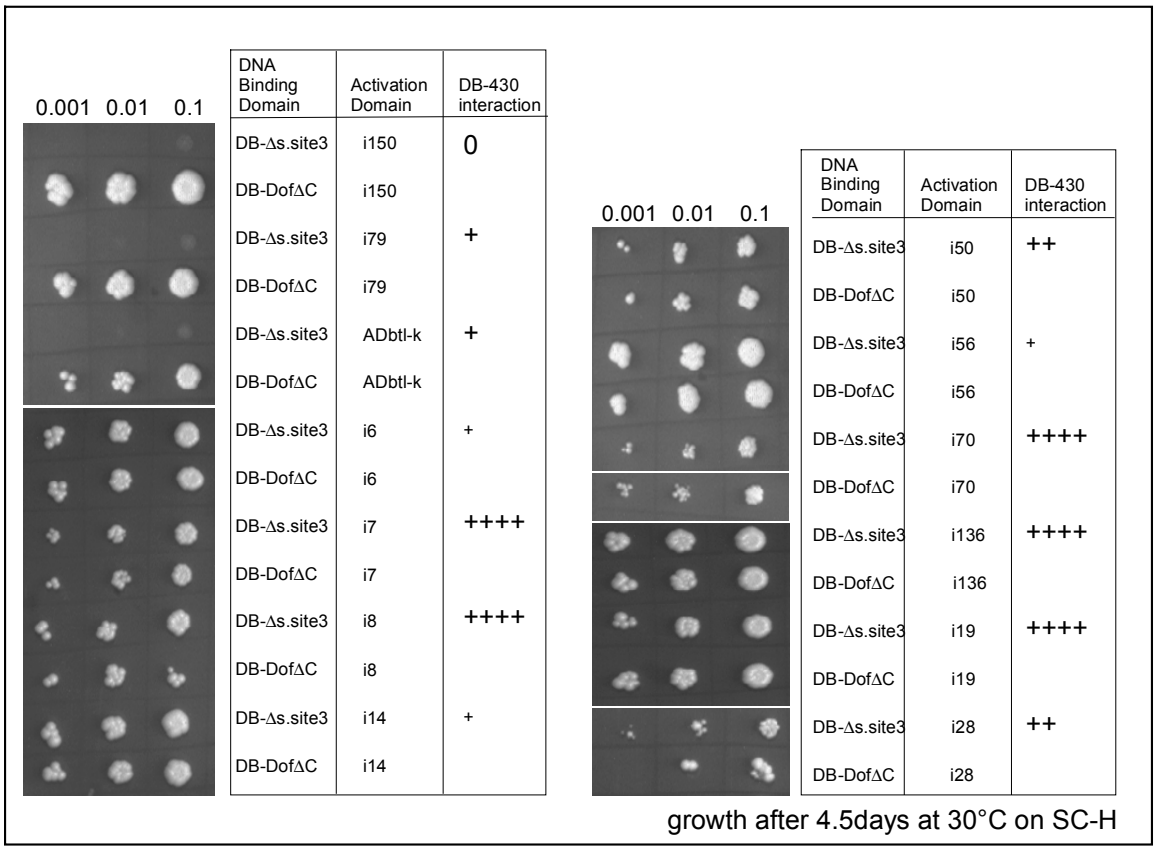


Figure 2.16 **B** shows the results for the activation of the reporter His3 by co-expression of the constructs shown in **A**. +++ = same growth as AD-DofΔC + DB-DofΔC co-expression after 4.5 days at 30°C on SC-H selection plates.

When co-expressed, DB–Δs.site3 and AD-Δs.site3 still activated the His3 reporter (Figure 2.16), indicating that if the third putative SUMOlation site were SUMOlated, it would not be the cause of the interaction between Hd#1. Surprisingly, the Δs.site3 construct did not interact with the Heartless receptor construct i150. When DB–Δs.site3 was co-expressed with the other Ade+ candidates, interaction of the His3 reporter was not

compromised for any of them, other than for the two FGFRs (Figure 2.17). It should be noted that while the interaction of DofΔC and i56 (the transcript coding for the AD-Ubc9 fusion protein) is diminished (Figure 2.17), this is not due to the mutation in the putative Lysine acceptor site, since Δs.site3 still binds strongly to i56.



DNA Binding Domain	Activation Domain	DB-430 interaction
DB-Δs.site3	i150	0
DB-DofΔC	i150	
DB-Δs.site3	i79	+
DB-DofΔC	i79	
DB-Δs.site3	ADbtl-k	+
DB-DofΔC	ADbtl-k	
DB-Δs.site3	i6	+
DB-DofΔC	i6	
DB-Δs.site3	i7	++++
DB-DofΔC	i7	
DB-Δs.site3	i8	++++
DB-DofΔC	i8	
DB-Δs.site3	i14	+
DB-DofΔC	i14	

DNA Binding Domain	Activation Domain	DB-430 interaction
DB-Δs.site3	i50	++
DB-DofΔC	i50	
DB-Δs.site3	i56	+
DB-DofΔC	i56	
DB-Δs.site3	i70	++++
DB-DofΔC	i70	
DB-Δs.site3	i136	++++
DB-DofΔC	i136	
DB-Δs.site3	i19	++++
DB-DofΔC	i19	
DB-Δs.site3	i28	++
DB-DofΔC	i28	

Figure 2.17 This figure shows the co-expression of the Ade+ candidates with either DB-DofΔC or DB-Δs.site3 on SC-H selection plates. For comparison, co-expression data of DB-430 (depicted in Figure 2.16) with the Ade+ candidates (Appendix II.) was included in the right column of the table. See Figure 2.3 for description of dot assay.

Expression pattern of *DmUbc9* and analysis of *DmUbc9* mutants

The i56 transcript, visualized by a Dig-labeled probe, is found ubiquitously at early stages but by stage 17 is restricted to the nervous system (Figure 2.17). The cytological location of *DmUbc9* is at 21C5-6. The *DmUbc9* mutant phenotype has been described (Epps and Tanda 1998) and has been shown to block nuclear import of the transcription factor Bicoid, resulting in multiple defects in anterior segmentation. The development of the tracheal system in embryos which lack *DmUbc9* (listed in Table 2.6) was examined by staining the tracheal lumen with the antibody mAB2A12. Stage 15-16 embryos were evaluated. No gross defects in tracheal development were observed. Thus, the dorsal trunk was fused, the dorsal, visceral and ganglionic branches were normal present. Embryos carrying the p-element insertion lwr^{02858} were stained with anti-Evenskipped to mark a subset of heart precursor cells which are absent in *dof* mutant embryos. Embryos

from P{ry$^{+t7.2}$=PZ}lwr^{02858} cn^{1}/CyO;ry^{506} parents showed a normal pattern of Evenskipped expressing cells. It is still possible that a maternal contribution of DmUbc9 is present in the *DmUbc9* homozygous mutant embryos and that defects in tracheal development or mesoderm migration could therefore not be determined.

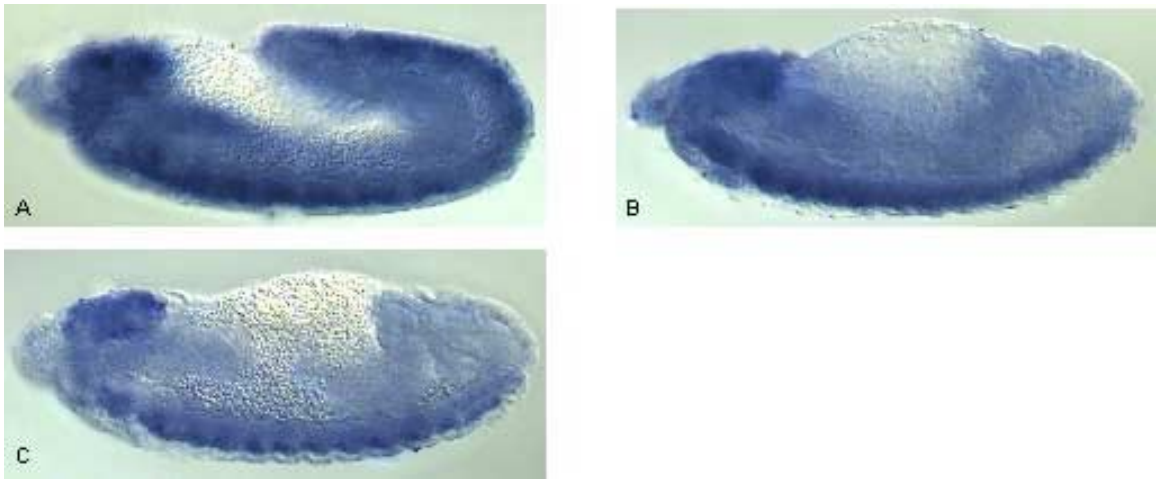


Figure 2.17 The expression pattern of *DmUbc9* (visualized with Dig-labeled i56 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 9 (A), 14 (B) and 17 (C). Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from the lateral view. Anterior is to the left, posterior to the right.

B#	genotype	description of mutagen
11214	P{ry$^{+t7.2}$=PZ}lwr^{02858} cn^{1}/CyO;ry^{506}	P-element P{PZ} inserted 560 bp downstream of putative polyadenylation signal for lwr (1)
11410	P{ry$^{+t7.2}$=PZ}lwr^{05486} cn^{1}/CyO;ry^{506}	P-element P{PZ} inserted into 5'UTR of lwr (1)
3548	Df(2L)al, ds^{al}/In(2L)Cy, Cy^{1}	Breakpoints: 021B08-C01;021C08-D01, 022D01-02;033F05-034A01 (2)
3084	Df(2L)ast2/SM1	Breakpoints: 021D01-02;022B02-03

Table 2.6 This table lists the mutant *DmUbc9 Drosophila* strains that were used to analyze embryos lacking *DmUbc9*. B#: Bloomington Stock Center stock number (1) Epps and Tanda 1998, Spradling et al. 1999 (2) Lindsley and Zimm

2.6 Further yeast two-hybrid interacting partners of Dof

As mentioned at the beginning of the results section, this chapter serves to document the investigations made with the other yeast two-hybrid partners of Dof.

(Ade+) Interacting partners implicated in cytoskeletal functions

Three candidates isolated in the yeast two hybrid screen are potentially mediators of cytoskeletal functions, namely βSpectrin, and Sinuous and the i19 group (i19). These will be presented in this sub-chapter.

βSpectrin

Transcripts corresponding to βSpectrin were isolated four times in the yeast two-hybrid screen. Sequence analysis of the 5 prime end shows that the longest transcript, i6, contains the last three spectrin repeats and the PH domain of beta-spectrin. The other three clones are just one amino acid shorter than i6. All four clones code for an in-frame fusion protein. βSpectrin is a component of the spectrin-based membrane skeleton, which confers mechanical stability to the membrane by being anchored to transmembrane proteins and to the actin cytoskeleton (Bray, 2000). βSpectrin associates with αspectrin to form a heterotetramer.

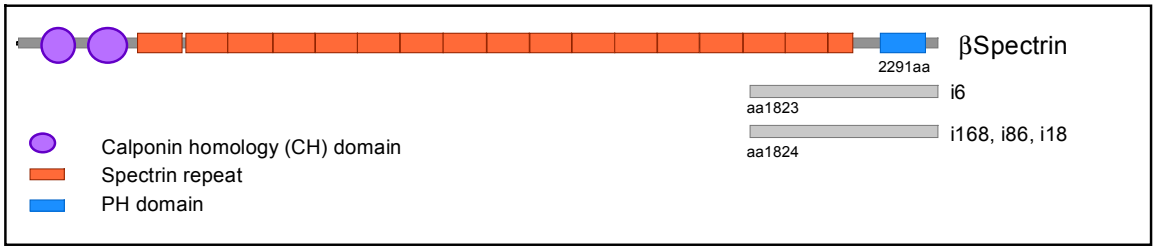


Figure 2.18 This figure shows the full length protein βSpectrin (sequence taken from the accession number AAA28399.1) aligned with the predicted amino acid sequences of the transcripts isolated in the yeast two-hybrid screen. The motifs shown in βSpectrin are names of PFAM motifs (see Materials and Methods) scoring higher than gathering threshold.

Expression Pattern of *βSpectrin* and analysis of *βSpectrin* mutants

Figure 2.19 shows embryos stained with Dig-labeled βSpectrin. The βSpectrin transcript is expressed ubiquitously until about stage 14. Then, high levels of transcript are found in the visceral mesoderm (A and B) , the central nervous system (C and D), the hindgut (F) and the posterior spiracles (E).

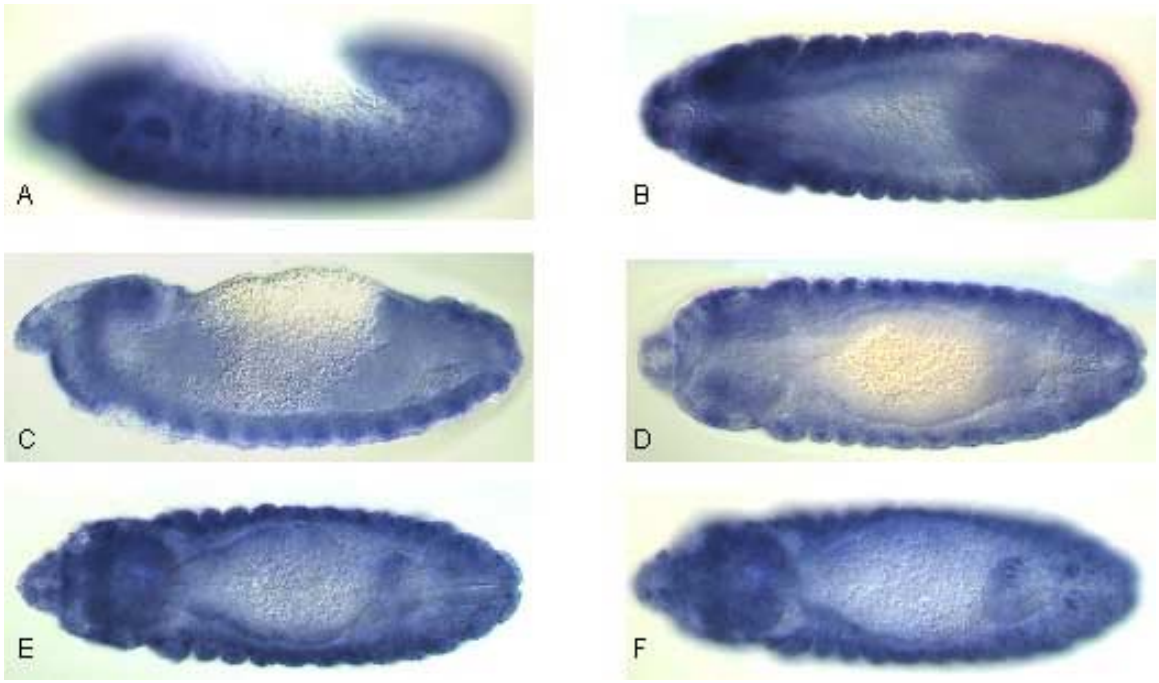


Figure 2.19 The expression pattern of βSpectrin (visualized with Dig-labeled i6 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 11 (A,B), 14 (C-F). Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Anterior is to the left, posterior to the right, A and C are lateral views, B,D,E and F are dorsal views.

The cytological location of βSpectrin is16C1. Table 2.7 shows the mutant strains used to analyze mesoderm and tracheal development in embryos lacking *βSpectrin*. Embryos were stained with the antibody mAB2A12 which marks the tracheal lumen and with anti-Evenskipped to mark a subset of heart precursor cells which are absent in *dof* mutant embryos. Stage 15-16 embryos were evaluated. No gross defects in tracheal development were observed and all three strains showed a normal pattern of *even skipped* expressing cells. Thus, if βSpectrin is required for Dof function, then the maternal contribution of βSpectrin is compensating for the lack of a zygotic transcript, or it has a redundant function, or it is required in other tissues in which Dof functions but that weren't analyzed here.

BL#	genotype	description of mutagen
4953	Df(1)BD10, r* f1/DP91;Y)W73/C(1)DX	breakpoints: 16A2;16C7-10,15C1-D6;16F;Yh1-h26
	Beta Spectrin em21/FM7 GFP	Ethyl-Methyl-Suffocate (1)
	Beta Spectrin em6/FM7 GFP	Ethyl-Methyl-Suffocate (1)

Table 2.7 This table lists the mutant *βSpectrin Drosophila* strains that were used to analyze embryos lacking *βSpectrin*. B#: Bloomington Stock Center stock number (1) Gift from R. Dubreuil

Sinous

The transcript i8 was isolated once in the screen and the 5 prime end is identical to CG10624, an annotated sequence on the third chromosome (64D3-4) showing no significant homology to any known gene. A P-element in the strain l(3)06524 is inserted at position 301 of the CG10624 open reading frame. The mutation arising from the insertion of the P-element l(3)06524 has been called "*sinuous*" and was identified in a screen for tracheal tube expansion mutants (Beitel et al. 2000). The insertion affects both tube length and tube diameter, resulting in enlarged and convoluted tracheal tubes (the dorsal trunk follows a tortuous path because it is too large) that are often constricted at fusion points.

The transcript of i8 is not in frame with the Gal4 activation domain. A new construct was made in which the reading frame was adjusted (construct 308) to code for the same amino acids encoded by the open reading frame of CG10624. This construct, however, did not activate the reporter His3 when co-expressed with DofΔC. This result will be discussed with the other candidates from the screen that were out of frame with the activation domain (see Discussion)

An alignment of the predicted amino acid sequence of i8 with that of CG10624 is shown in Figure 2.20. The amino acid sequence of an in-frame i8 transcript is 19 amino acids short of the predicted protein sequence of CG10624. The 3' end of i8 contains an exon not included in CG10624, but the additional sequence shows no homology to anything in the database and no motif for the translated sequence was found in PFAM. A PFAM motif search with CG10624 revealed a similarity (albeit below gathering threshold) to the PMP22_Claudin domain. Proteins with the PFAM "PMP22-Claudin" domain (accession number PF00822) are integral membrane proteins of about 200 residues. They have four conserved transmembrane regions and are localized to tight junctions, the vertebrate equivalent of *Drosophila* septate junctions. A search for transmembrane regions (TM PRED, ISREC Server) in the predicted protein sequence of CG10624 predicts 4 transmembrane regions having a strongly preferred intracellular N-terminus.

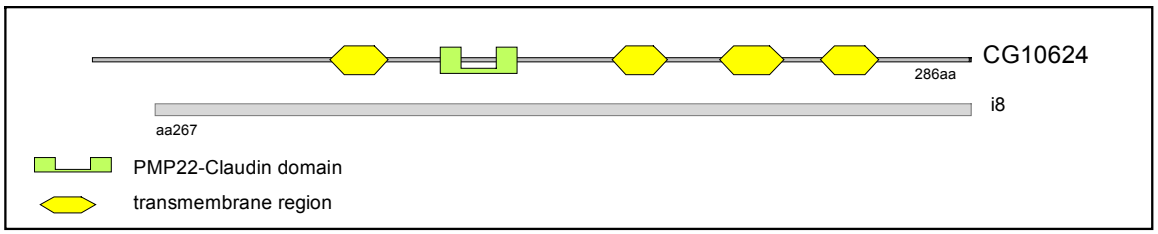


Figure 2.20 This figure shows the predicted amino acid sequence for the predicted gene CG10624, also called *Sinuous* (sequence taken from the open reading frame of CG10624) aligned with the predicted amino acid sequence of i8 (using the same reading frame as that of CG10624). The PMP22-Claudin domain shown in Sinuous is a PFAM motifs (see Materials and Methods) scoring lower than gathering threshold. The transmembrane regions were predicted by the program TM PRED

The *in situ* hybridization with Dig-labeled antisense i8 RNA in Figure 2.21 shows that sinuous is expressed in the ectoderm (A,B) in the hindgut (C), but is absent from the ventral ectoderm. The ectodermal expression is not uniform. Higher expression levels can be seen in a segmental pattern (A, B, E). At later stages, the transcript is found in the

dorsal branch, but this might be background staining, since the staining is found in the lumen and not in the cells.

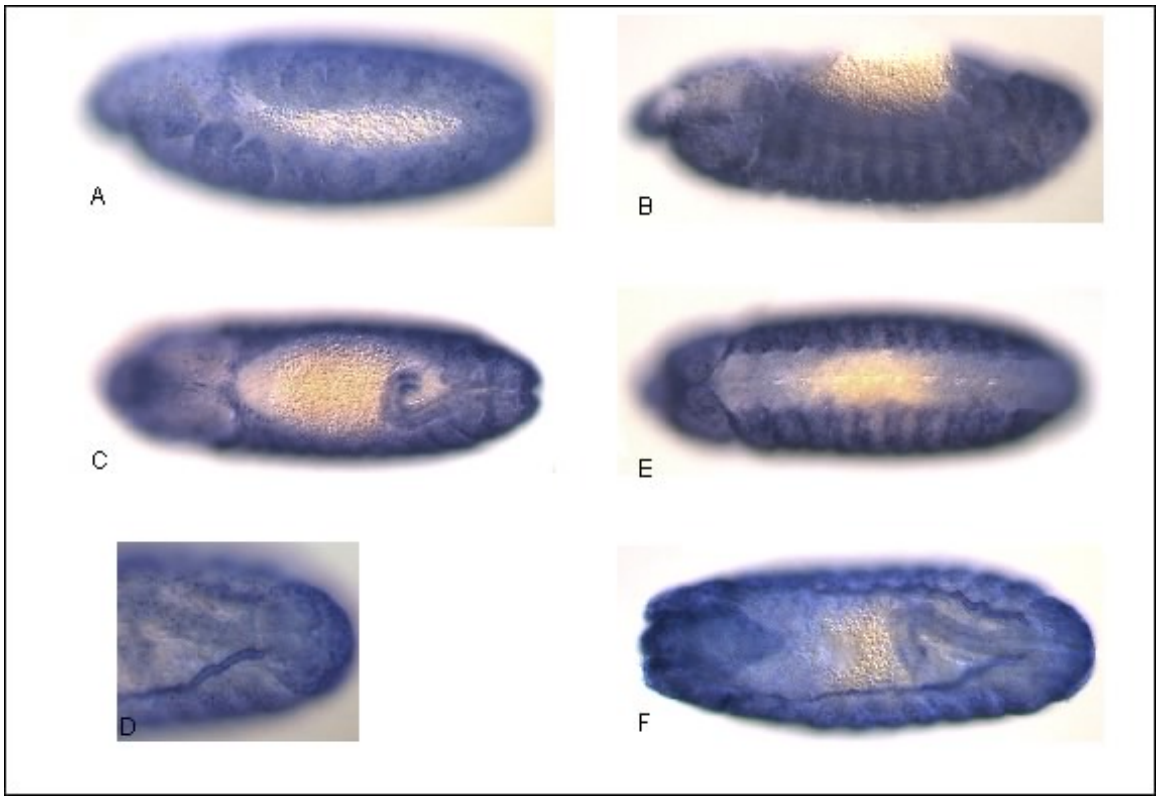


Figure 2.21 This figure shows the expression pattern of *sinuous* (visualized with Dig-labeled i8 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 9 (A), 14 (B-E) and 17 (D,F). Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Anterior is to the left, posterior to the right, A and B are lateral views, C,D and F are dorsal views, E is a ventral view. D is a higher magnification of the posterior dorsal trunk in F.

i19

The group i19 consists of three members. All three clones code for an in-frame fusion protein. Part of the 5 prime end of the three clones matches the predicted gene CG2008 whose protein sequence contains two PH domains and a Dbl homology (DH) domain (Figure 2.22). As can be seen in Figure 2.22, additional 5 prime sequence is contained in the i19 transcripts, suggesting that the annotated gene CG2008 does not reflect the entire gene coding for the i19 group transcripts. Further sequence analysis of the longest clone, i19, showed that the 3 prime and 5 prime sequence of i19 is contained in the expressed sequence tag (EST) SD10794, whose sequence encompasses two annotated sequences, CG2008 and CG1288. The protein sequence of CG1283 contains a FERM domain (4.1 – ezrin – radixin – moesin domain). The clone SD10794 was sequence and compared to other available sequences in this region. The results are summarized in Figure 2.23 and the predicted borders of each exon are listed in Table 2.7.

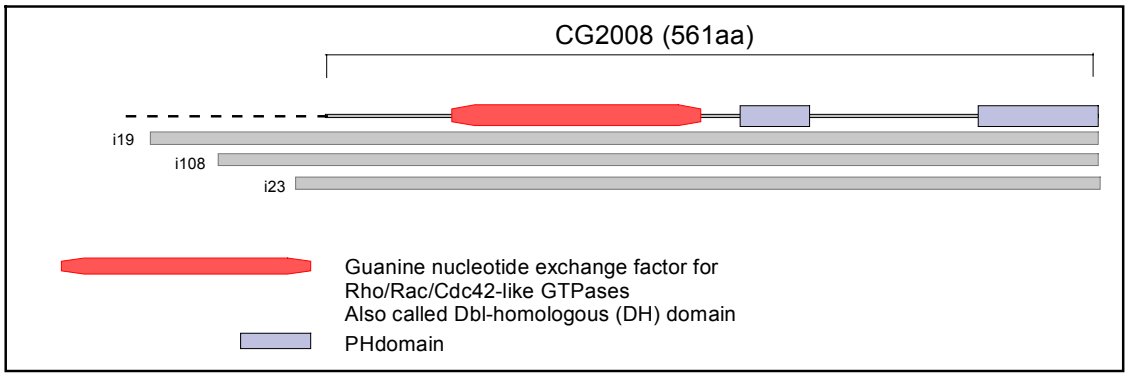


Figure 2.22 This figure shows the full length protein encoded by CG2008 aligned with the predicted amino acid sequence of the transcripts of the i19 group. The motifs shown in CG2008 are names of PFAM motifs (see Materials and Methods) scoring higher than gathering threshold. Note that all three clones contain additional sequence at the 5' end.

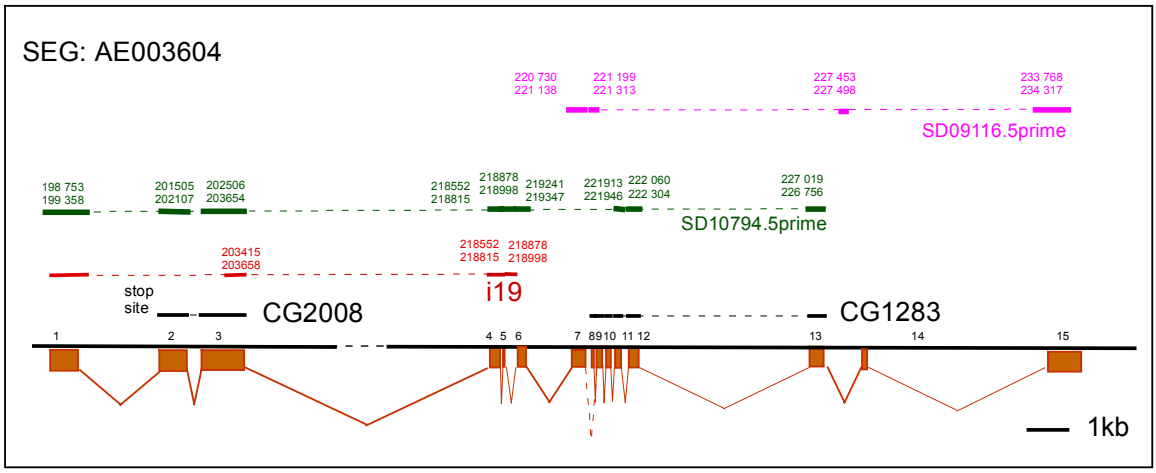


Figure 2.23 This figure is an alignment of the sequence of SD09116 and SD10794, CG1283, CG2008 and i19 on the genomic sequence SEG: AE003604. The predicted exons are shown as brown boxes. For each sequence analyzed, dashed lines indicate regions that weren't sequenced, whereas filled lines indicate regions for which sequence data is available. Numbers on top of each filled line indicate the approximate 5 prime and 3 prime position on the genomic sequence to which the sequence matches.

exon	5 prime	3 prime	sequence data
1	199358	199876	i19, SD10794
2	201 505	202 107	CG2008, SD10794
3	202506	203 654	CG2008, i19, SD10794
4	218 552	218 814	i19, SD10794
5	218 878	218 998	i19, SD10794
6	219241	219347	SD10794
7	220 730	221 138	SD09116.3prime
8	221 038	221 145	CG1283
9	221 192	221 362	CG1283, SD09116.3prime (3' only)
10	221 411	221 663	CG1283,
11	221 713	221 946	CG1283, SD10794.5prime (5' only)
12	222 060	666 304	CG1283, SD10794.5prime
13	226 756	227 019	CG1283, SD10794.5prime
14	227 453	227498	SD09116.5prime (5' only)
15	233 768	234 317	SD09116.5prime

Table 2.7 lists the approximate beginning and end of each putative exon in the region AE003604 and the sequence data used to map these positions.

Motifs in SD10794 and its homology to CDEP

The sequence analysis of SD10794 shows that it encodes a protein with a FERM domain, a Dbl homology (DH) domain and two PH domains. When the sequence of SD10794 is compared to other sequences in the database (search method used: TBLASTX), it shows high homology (E-value e-174) to human CDEP (chondrocyte derived ezrin protein). CDEP was isolated using the subtractive hybridization method between dedifferentiated cartilage cells and overtly differentiated cartilage cells (Koyano et al. 1997). As suggested by the high homology to SD10794, CDEP also contains a FERM domain, a PH domain and a DH domain. All three domains are implicated in cytoskeletal functions. FERM domains are found in a number of cytoskeletal-associated proteins that associate with various proteins at the interface between the plasma membrane and the cytoskeleton (Chishti AH et al. 1998). Rho guanine nucleotide exchange factors convert the inactive, GDP-bound Rho to active, GTP-bound Rho, and thus regulate the activity of the small GTPase Rho. Rho family GTPases are considered to be a major regulator of actin mediated cytoskeletal changes (reviewed in Hall and Bar-Sagi, 2000). PH domains, as discussed in the Introduction, potentially mediate membrane localization of proteins in response to signaling.

i19 expression pattern and analysis of Deficiencies

At stage 10, the i19 transcript is expressed in the somatic and head mesoderm, as well as at the hindgut invagination. At stage 14, the transcript is found in many tissues, including the visceral mesoderm, the head mesoderm, the hindgut and, possibly the tracheae.

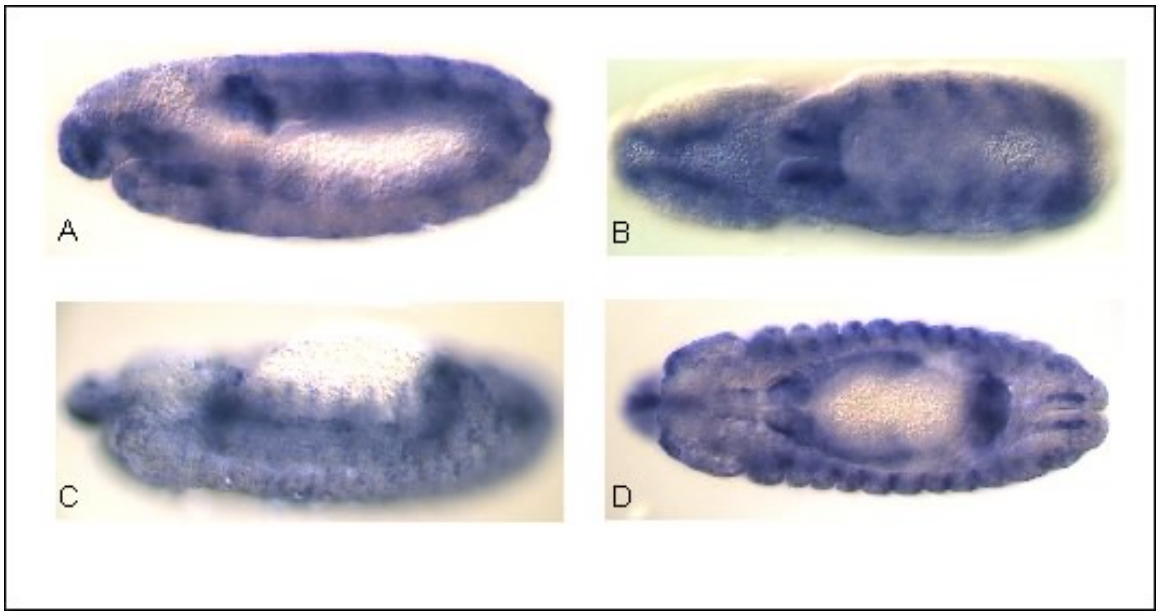


Figure 2.23 This figure show the expression pattern of i19 (visualized with Dig-labeled i19 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 9 (A, B), and 14 (C,D) Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from a lateral view in A and C and from a dorsal view in B and D. Anterior is to the left, posterior to the right.

One way to establish whether gene i19 is involved in the same signaling cascade as Dof is to determine whether i19 is also required in tracheal morphogenesis or mesoderm spreading by analyzing an embryo that lacks i19. Ideally, the gene would be disrupted by a point mutation or a transposon insertion, but no such mutation has been recorded in Flybase. The i19 transcript locates to position 82E on the cytological map. The closest mapped P-element insertions are in the gene *corto* at 82F5 (EP(3)0775) and in CG14656 at 82E1 (l(3)02733 and EP(3)3416). The relatively small deficiency, Df(3R)107, removes the genomic region at 82E01-3. This deficiency has been shown by complementation analysis to delete the genes *opa*, *Gnf1*, *l(3)82Ea* and *l(3)82Eb* (Flybase, Megraw and Cook, 1998.10.2) but does not delete *corto* or *l(3)02733*. Using the genomic sequence and the aforementioned complementation analysis, it is possible to limit the deleted region of Df(3R)107 down to a maximum of 18 annotated genes (including CG2008 and CG1283), two of which have been described, namely *opa* and *Gnf1*.

Although Df(3R)107 removes i19, it is not possible to use this deficiency strain to analyze the phenotype of a loss of i19 since the gene *opa*, which is also removed, disrupts the orderly development of segment polarity, thus obfuscating any phenotype caused by the loss of i19. l(3)82Ea and l(3)82Eb (Flybase, Megraw and Cook, 1998.10.2) are two strains containing lethal mutations which map to this region. Since the two lethal mutations map to the same region as i19, there was a possibility that these mutations disrupt i19. The mutant embryos from heterozyogous l(3)82Ea and l(3)82Eb parents were examined by staining the tracheal lumen with the antibody mAB2A12, marking a subset of heart precursor cells with anti-Evenskipped and marking the visceral mesoderm with anti-FasciclinIII. Neither mutant allele showed an abnormal phenotype. The deficiencies and lethal alleles that were analyzed are listed in Table 2.8.

BL#	Genotype	comments
BL76	w^{*}; Df(3R)107, e^{*}/TM3, P{$w^{+mW.hs}$=Thb8-lacZ}WD1, Sb^{1} Ser^{1} breakpts082E01-03;082E01-03	pair-rule segmentation defects (due to *opa*)
BL4841	w^{1118}; l(3)82Ea^{T6}/TM6B, Tb^{1}	wild type trachea
BL4810	l(3)82Ea^{l18}/TM3, Sb^{1}	wild type trachea after rebalancing
BL4811	l(3)82Eb^{T10} e^{T10}/TM6B, Tb^{1}	wild type trachea
BL4812	l(3)82Eb^{T47}/TM3, Sb^{1}	messy embryos possibly due to balancer

Table 2.8 This table lists the mutant *Drosophila* strains that were used to analyze embryos lacking the gene coding for i19. B#: Bloomington Stock Center stock number

In summary, the genetic analysis of embryos lacking the gene coding for i19 was not completed. All deficiency chromosomes in this region remove *odd paired*, and therefore cannot be used to determine whether the lack of i19 disrupts the same developmental processes in which Dof is involved. l(3)82Ea and l(3)82Eb are two lethal mutations which map to this region (complementation analysis). Since the two lethal mutations map to the same deficiency as i19, it was possible that these mutations disrupt i19. However, neither mutant strain showed any abnormal phenotype in the tissues assayed. This leaves the following possibilities if i19 were to affect the same functions as Dof: the two lethal strains are lethal mutations in genes different from i19, the maternal component might compensate for the lack of zygotic i19 and finally, i19 might affect other morphogenetic process in which Dof is involved but were not assayed here.

Other Ade+ Interacting partners

In this section, all interaction partners that are not described above and that activated the Ade2 reporter when co-expressed with DofΔC are described here. The clone i136 (whose 5 prime sequence matches that of the predicted gene CG10671) is not described here since it was not analyzed further.

i7

The group i7 consists of two clones. The transcript maps to SEG: AE003433 and maps cytologically to 4D2-F2. Sequence analysis of the 5' end of transcripts i7 and i104 shows no homology to any known expressed sequence tag or annotated sequence, nor does the predicted protein sequence have any PFAM motifs. No mutations in this gene have been recorded and no embryos in which this gene is removed were analyzed. The expression pattern of the gene coding for i7 is shown in Figure 2.24. The i7 transcript can be detected in the ectoderm at early stages and in the central nervous system and tracheal lumen at later stages (C,D,E,F,G).

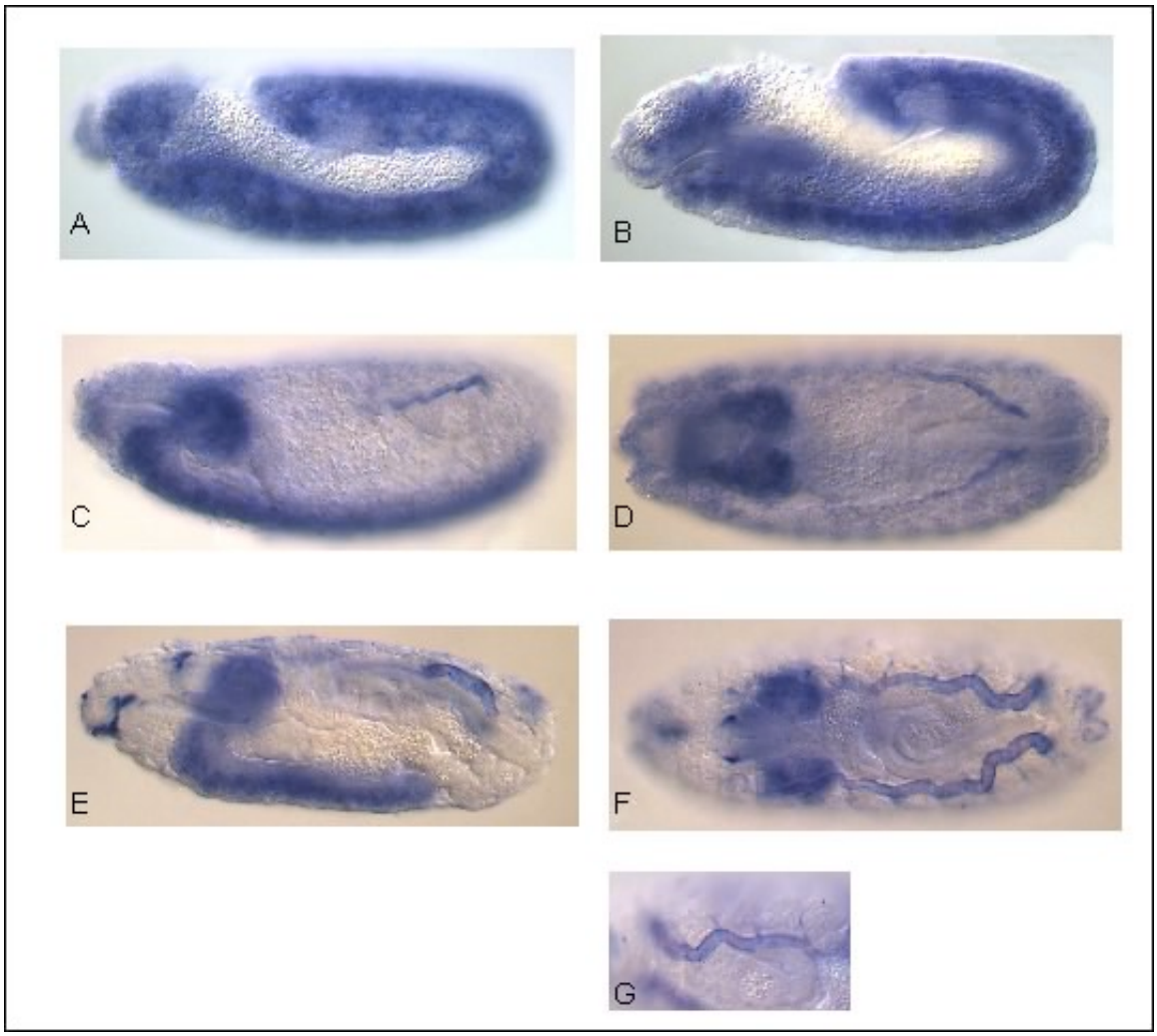


Figure 2.23 This figure show the expression pattern of i7 (visualized with Dig-labeled i7 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 7 (B), 9 (A) 16 (C,D) and 17 (E,F,G) Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from a lateral view in A,B,C and E and from a dorsal view in D and F. G is a higher magnification of the dorsal branch in F. Anterior is to the left, posterior to the right.

i159

The group i159 consists of three identical clones. The transcript maps to SEG: AE003492 and maps cytologically to 11F-12A. Sequence analysis of the 5' end of transcripts i159 shows no homology to any known expressed sequence tag or annotate genes, nor does the predicted protein sequence have any PFAM motifs. The expression pattern of the gene coding for i159 is shown in Figure 2.24 and shows that it is expressed ubiquitously at early stages but becomes restricted to the central nervous system at later stages.

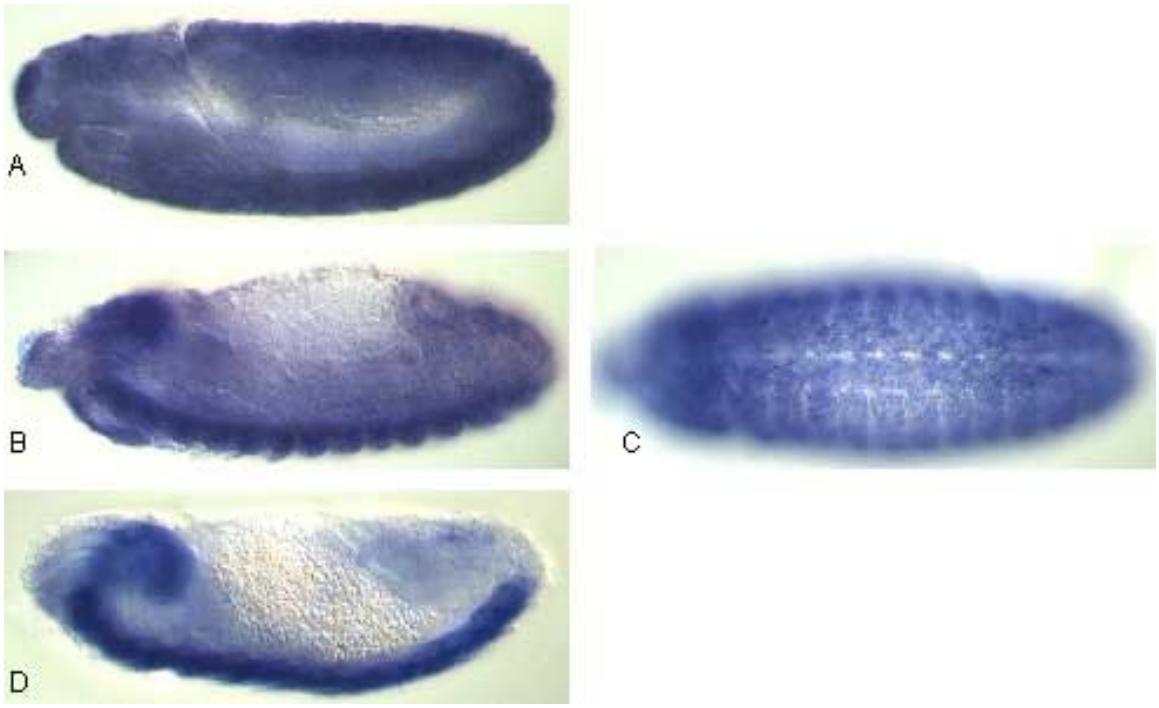


Figure 2.24 This figure show the expression pattern of i159 (visualized with Dig-labeled i159 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 11 (A), 14 (B,C) and 15 (D) Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from a lateral view in A,B and D and from a ventral view in C. Anterior is to the left, posterior to the right.

Embryos deficient for the genomic region 11F-12A from parents of the strains listed in Table 2.9 were examined by staining the tracheal lumen with the antibody mAB2A12 and marking a subset of heart precursor cells with anti-Evenskipped. All embryos had wild

type trachea and normal *Evenskipped* expressing heart precursor cells. The deficiencies and lethal alleles that were analyzed are listed in Table 2.8.

BL#	genotype	comments
966	Df(1)N12, $ras^{1}v^{1}$/FM6	breakpoints 11D1-2;11F1-2
967	Df(1)C246/FM6	11D-E;12A1-2

Table 2.8 This table lists the mutant *Drosophila* strains that were used to analyze embryos lacking the genomic region 11F – 12A. B#: Bloomington Stock Center stock number

Haywire

The transcript i50 was isolated in the yeast two-hybrid screen as a single clone. Sequence analysis of the 5 prime end of the clone showed that the transcript is from the gene *haywire*, which encodes a general RNA polymerase II transcription factor (Figure 2.25). The haywire cDNA in construct i50 is not in frame with the Gal4 activation domain.

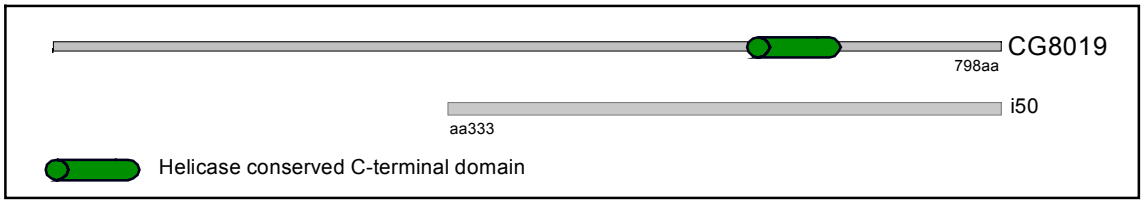


Figure 2.25 This figure shows the predicted amino acid sequence encoded by *haywire* (CG8019) aligned with the predicted amino acid sequence of i50 (using the same reading frame as that of *haywire*). The motifs shown in CG8019 are names of PFAM motifs (see Materials and Methods) scoring higher than gathering threshold

The expression pattern of *haywire* is shown in Figure 2.24 and shows that it is expressed ubiquitously at all stages.

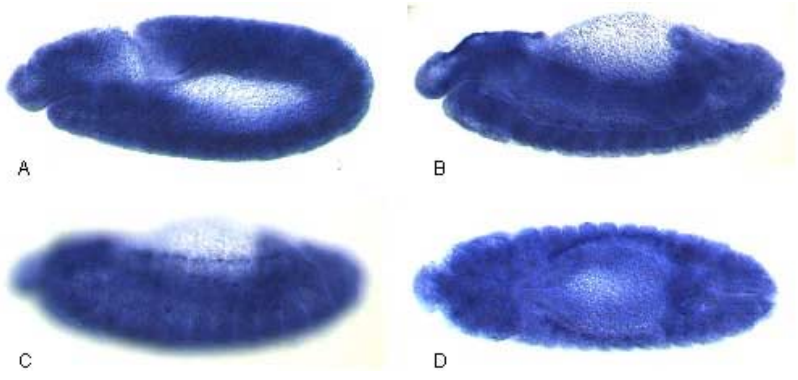


Figure 2.26 This figure show the expression pattern of i50 (visualized with Dig-labeled i150 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 11 (A,B) and 14 (C,D).

Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from a lateral view in A,B and C and from a dorsal view in D. Anterior is to the left, posterior to the right.

Mutations in *haywire* exist and have been described. For the analysis here the deficiency Df(3L)lxd6, in which *haywire* is deleted, was used to determine whether the lack of haywire might affect tracheal development or mesoderm migration. The deficiency y[1?];Df(3L)lxd6/TM3, Sb[1] Ser[1] has its breakpoints between 067E01-02 and 068C01-02. As can be seen if Figure 2.27, the tracheae do not develop as in wild type embryos: all branches develop except for the dorsal branch. The subset of heart precursor cells which do not develop in *dof* mutant embryos do develop normally in these embryos.

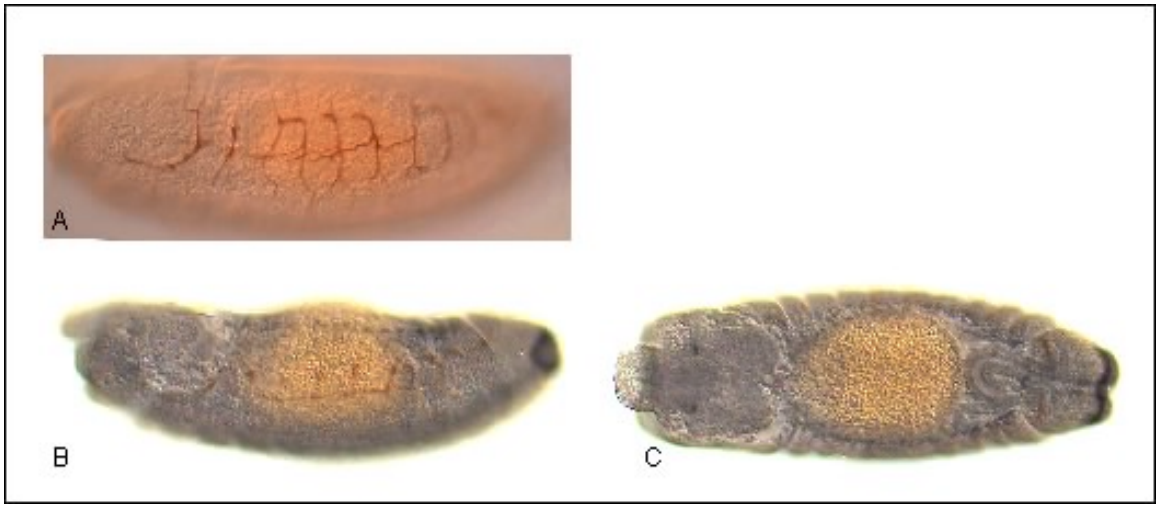


Figure 2.27 This figure shows an embryo from parents of the genotype y[1?];Df(3L)lxd6/TM3, Sb[1] Ser[1]. The embryo in A is stained with the antibody mAB2A12, which marks the tracheal lumen. Embryos in B and C are stained with mBA2A12 (in brown) and anti-Evenskipped (in blue), which marks a subset of heart precursor cells. Embryos are shown from a lateral view in A and B and from a dorsal view in C. Anterior is to the left and posterior is to the right.

Hr39

The transcript i70 was isolated once in the yeast two-hybrid screen. DNA sequence analysis of the 5 prime end showed that i70 is a transcript of *Hormone receptor-like in 39* (*Hr39*), also known as FTZ-F1β, which encodes an orphan member of the nuclear receptor superfamily (Ayer et al. 1993; Ohno and Petkovich 1993). Hr39 is synonymous with CG8676 and its gene product codes for a nuclear receptor. The transcript i70 encodes an in-frame fusion with the Gal4 activation domain and the i70 transcript codes for a protein that starts at the equivalent of amino acid 446 of *Hr39*. An *in situ* hybridization with dig-labeled i70 shows that Hr39 is expressed in the ectoderm (Figure 2.28).

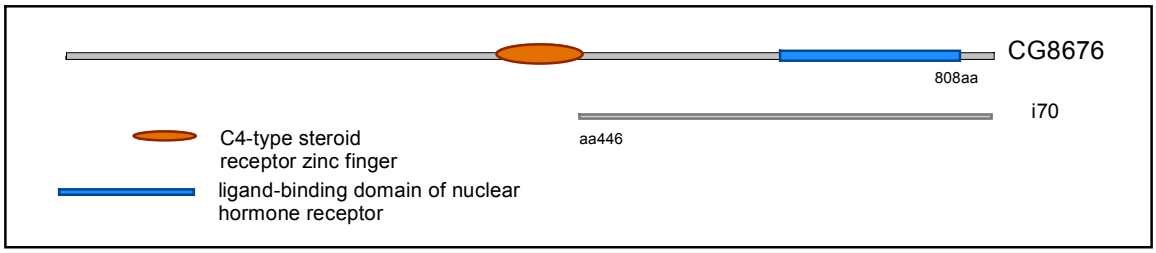


Figure 2.27 This figure shows the predicted amino acid sequence encoded by *Hr39* (CG8676) aligned with the predicted amino acid sequence of i70. The motifs shown in CG8676 are names of PFAM motifs (see Materials and Methods) scoring higher than gathering threshold

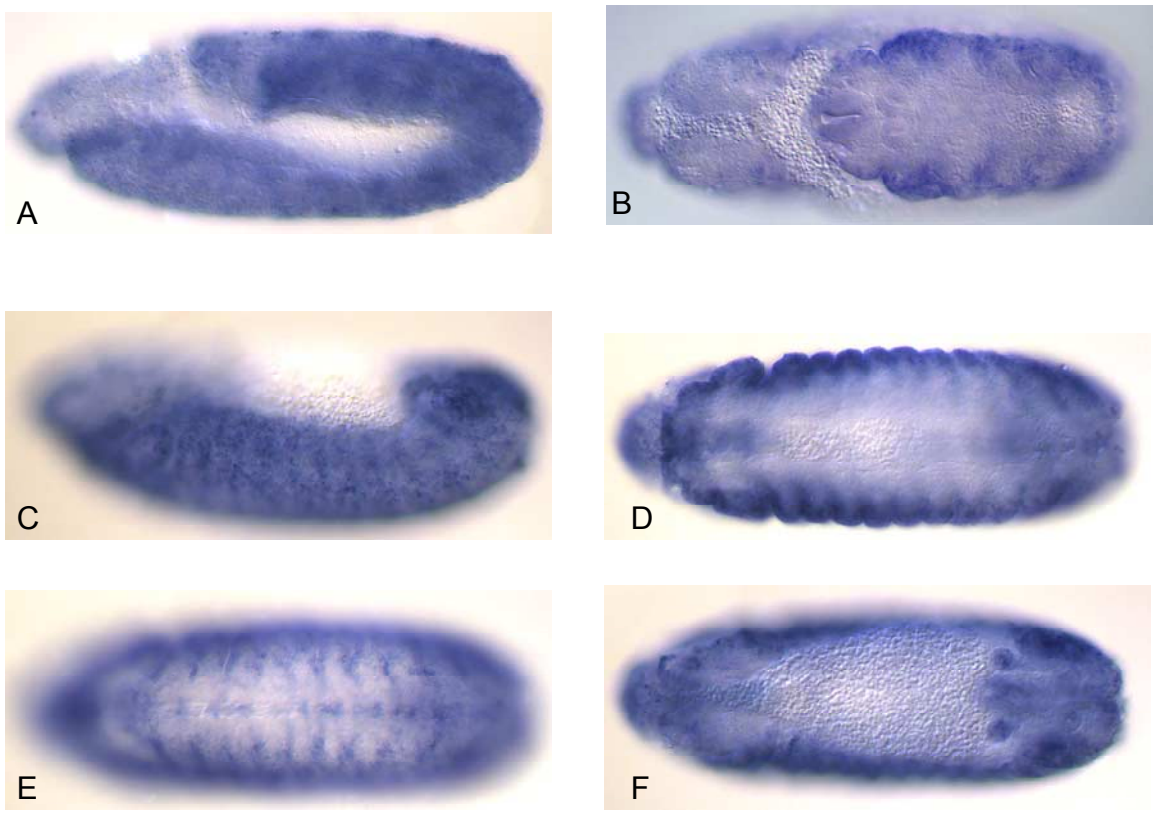


Figure 2.28 This figure show the expression pattern of i70 (visualized with Dig-labeled i70 anti-sense RNA) in wild type embryos that are approximately at the developmental stage 10 (A, B), and 14 (C-F) Staging according to Campos-Ortega, J.A.; Hartenstein, V. 1997. Embryos are shown from a lateral view in A and C and from a dorsal view in B and F and from a ventral view in E and D. Anterior is to the left, posterior to the right.

The cytological position of *Hr39* is 39C1-4. A P-element insertion in *Hr39* exists and is called $Hr39^{k13215}$. The P-element insertion maps to the first intron but does not affect

viability. Whether the P-element insertion generates a null allele of Hr39 is unclear, since residual Hr39 expression could be detected in a Northern Blot of mRNA from $Hr39^{k13215}$ embryos (Thummel, D.I.S.80, 1997). Df(2L)TW1/CyO is a deficiency strain in which the region 038A07-B01 to 039C02-03 is deleted. This deficiency was initially used to determine whether the absence of i70 leads to a disruption in tracheal development. However, the deficiency removes the EGFR ligand Spitz, and EGF signaling is required during tracheogenesis (Wappner et al. 1997, Llimargas and Casanova 1999). The trachea of these embryos do not develop normally (Figure 2.29) and the phenotype of the deficiency is as strong as that of embryos that are hemizygous for Spi2A14 and Df(2L)TW1/Cyo, indicating that the severe phenotype found in Df(2L)TW1/Cyo embryos is due mainly to the absence of *Spitz*.

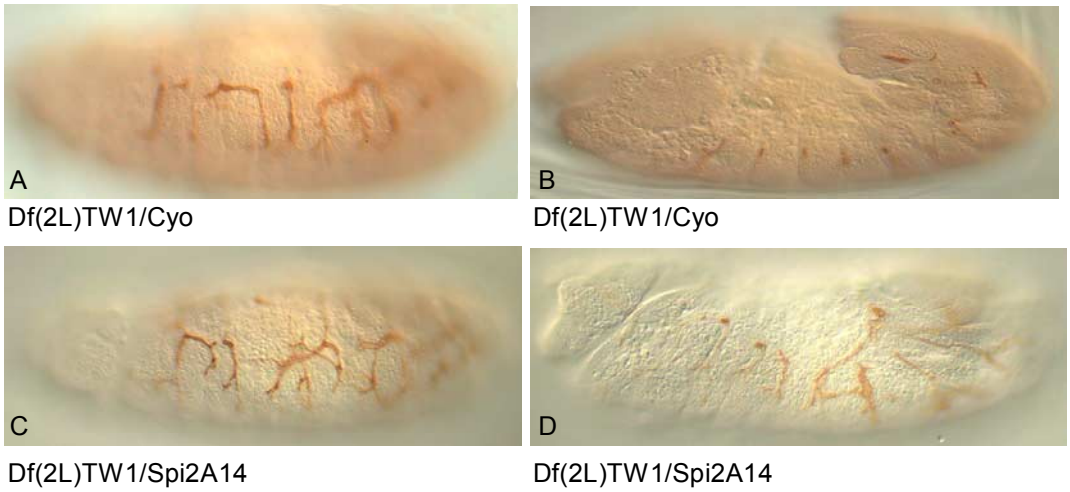


Figure 2.29 This figure shows embryos stained with mAB2A12, an antibody which marks the tracheal lumen. The two embryos Spi2A14 flies were a provided by S. Roth. The embryos in B and D are arrested in germ band extension and the embryos in A and C are at about stage 14.

i14

The interactor group i14 contains 10 members. Sequence analysis of the 5 prime end shows that the transcript derives from the gene predicted by CG14206. The amino acid sequence of CG14206 is homologous to human rpS10. The sequence analysis of the 5 prime end of i14 members revealed two further features. First of all, all the transcripts are out of frame. That is, were the reading frame of the Gal4 activation domain upheld, only a few codons would code for an amino acid, before the first stop is encountered. None of the interactors contain the reading frame that would give rise to the CG14206 protein, in fact, four clones are in the "+2" reading frame, and the other 6 are in the "+3" reading frame. The second observation is that all clones contain the 5' UTR. This is shown in Figure 2.30

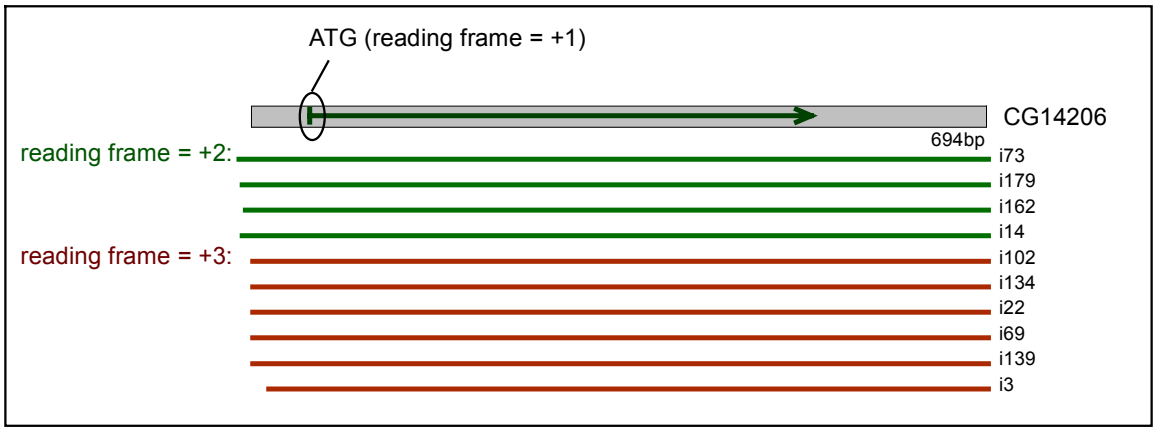


Figure 2.30 The DNA sequence of CG14206 is shown aligned to the transcripts from the i14 group. The open reading frame of CG14206 is shown by an arrow. Clones that are in the "+2" reading frame are shown in green and clones that are in the "+3" reading frame are shown in red.

The longest clone in the i14 group is i73 and its 5 prime end is identical to the expressed sequence tag cDNA LP04523. The sequence CG14206 maps to the position 18D7 on the first chromosome. There are no recorded deletions in this region. The sequence of CG14206 is highly homologous to CG12275. Which maps to the position 98B on the third chromosome. There are two P-element strains that have an insertion in CG12275, namely one at base 417 (EP(3)3203) and the other at base 249 (EP(3)3683). The trachea of embyros homozygous for each insertion were examined using the lumenal marker mAB2A12, but no obvious defects were observed.

Candidates that activate the His3, but not the Ade2 reporter

The candidates that did not activate the Ade2 reporter when co-expressed with DofΔC are listed in Figure 2.31. The Figure includes the deficiency strains analyzed to determine whether the lack of a certain gene disrupts tracheal development or mesoderm migration (see figure legend for more details). Appendix III reports the analysis of the group i119, which was analyzed in detail because the deficiency strain which removes this region showed a strong phenotype. However, when a mutation was made in the gene corresponding to the i119 group, the tracheal development was found to be normal.

Figure 2.31 This figure lists the clones that did not activate the Ade2 reporter when co-expressed with DofΔC. The columns contain the following information **1**. the position of the first base of the candidate clone that matches the full length cDNA (CG annotated genes were used for this figure). **2**. here it is noted whether the transcript is in-frame with the activation domain fusion protein. **3**. The length of the cDNA sequence used in column 4. **4**. A figure of the cDNA sequence (thin bar), the open reading frame (thick bar) and PFAM motifs (squares or rectangles). The sequence contained in the interacting candidates is shaded in grey and if more than one candidates belong to a group, the first base of the shortest clone is indicated by an arrow. **5**. This column gives the name of the gene and the CG annotated sequence used for this figure **6**. Cytology **7**. Bloomington number of Deficiency analyzed **8**. Genotype of Deficiency analyzed **9**. Comments about Deficiency analyzed.

Figure 2.31

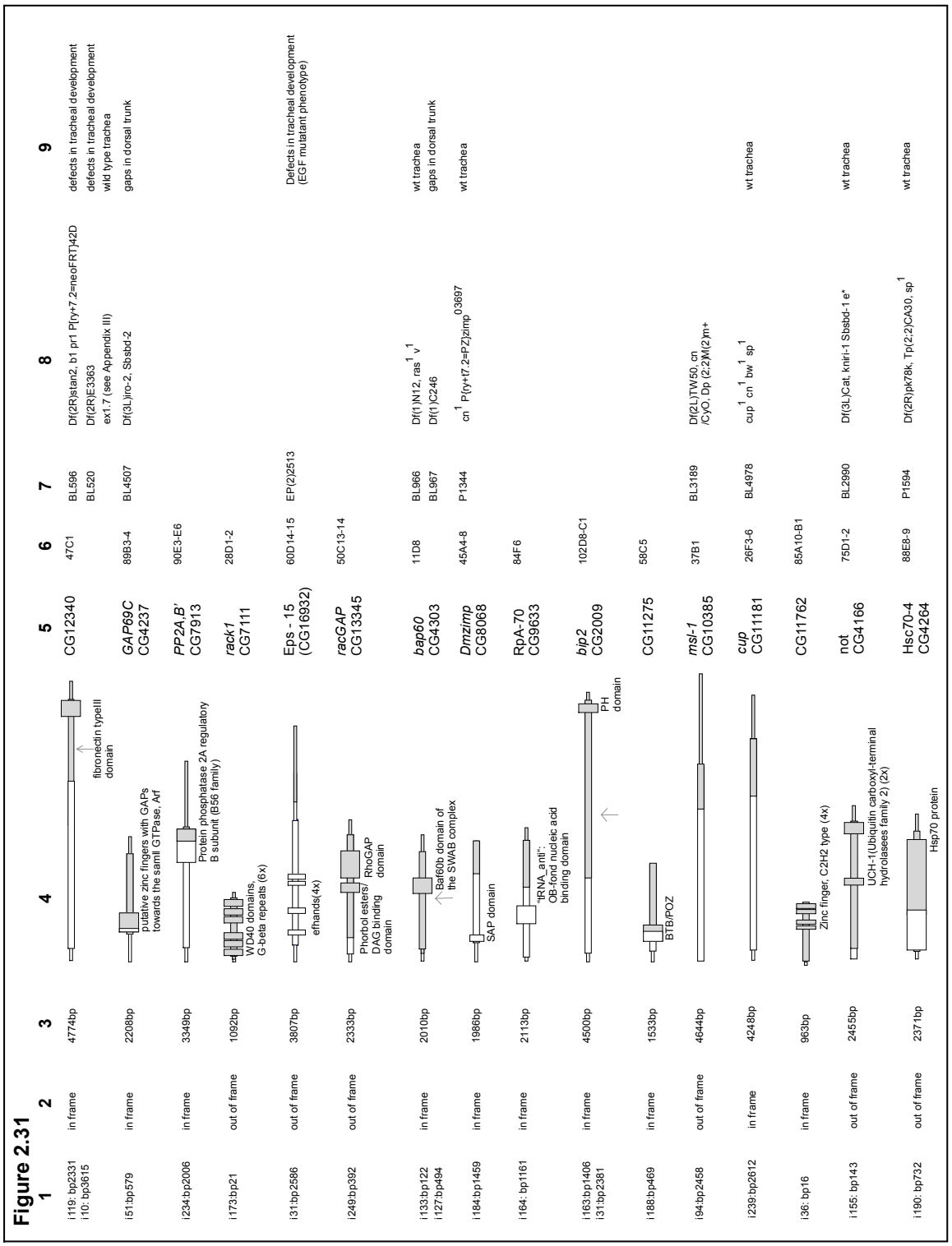
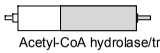
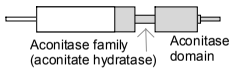
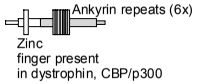
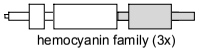
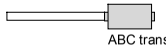

1
2
3
4
5
6
7
8
9
i119: bp2331
i10: bp3615
in frame
4774bp
fibronectin typeIII domain
CG12340
47C1
BL596
BL520
Df(2R)kstan2, b1 pr1 P{ry+7.2=neoFRT}42D
Df(2R)E3363
defects in tracheal development
defects in tracheal development
i51:bp579
in frame
2208bp
putative zinc fingers with GAPs towards the samll GTPase, Arf
GAP69C
CG4237
89B3-4
BL4507
ex1.7 (see Appendix III)
Df(3L)iro-2, Sbsbd-2
wld type trachea
gaps in dorsal trunk
i234:bp2006
in frame
3349bp
Protein phosphatase 2A regulatory B subunit (B56 family)
PP2A,B'
CG7913
90E3-E6
i173:bp21
out of frame
1092bp
WD40 domains, G-beta repeats (6x)
rack1
CG7111
28D1-2
i31:bp2586
out of frame
3807bp
efhands(4x)
Eps - 15
(CG16932)
60D14-15
EP(2)2513
Defects in tracheal development
(EGF mutatant phenotype)
i249:bp392
out of frame
2333bp
Phorbol esters/ DAG binding domain
RhoGAP domain
racGAP
CG13345
50C13-14
i133:bp122
i127:bp494
in frame
2010bp
Baf60b domain of the SWAB complex
bap60
CG4303
11D8
BL966
BL967
Df(1)N12, ras^{1} v^{1}
Df(1)C246
wt trachea
gaps in dorsal trunk
i184:bp1459
in frame
1986bp
SAP domain
Dmzimp
CG8068
45A4-8
P1344
cn^{1} P{ry+t7.2=PZ}zimp^{03697}
wt trachea
i164: bp1161
in frame
2113bp
"tRNA_anti": OB-fond nucleic acid binding domain
RpA-70
CG9633
84F6
i163:bp1406
i31:bp2381
in frame
4500bp
PH domain
bip2
CG2009
102D8-C1
i188:bp469
in frame
1533bp
BTB/POZ
CG11275
58C5
i94:bp2458
out of frame
4644bp
msl-1
CG10385
37B1
BL3189
Df(2L)TW50, cn /CyO, Dp (2;2)M(2)m+
i239:bp2612
in frame
4248bp
cup
CG11181
26F3-6
BL4978
cup^{1} cn^{1} bw^{1} sp^{1}
wt trachea
i36: bp16
in frame
963bp
Zinc finger, C2H2 type (4x)
CG11762
85A10-B1
i155: bp143
out of frame
2455bp
UCH-1(Ubiquitin carboxyl-terminal hydrolasees family 2) (2x)
not
CG4166
75D1-2
BL2990
Df(3L)Cat, kniri-1 Sbsbd-1 e^{s}
wt trachea
i190: bp732
out of frame
2371bp
Hsp70 protein
Hsc70-4
CG4264
88E8-9
P1594
Df(2R)pk78k, Tp(2;2)CA30, sp^{1}
wt trachea

Figure 2.31 (continued)

1	2	3	4	5	6	7	8	9
i207:bp1	out of frame	1591bp		CG14805	2B14			
i130:bp664	out of frame	1752bp		CG7920	99D3-5			
i59:bp1344 i204:bp1755	in frame	2631bp		aconitase CG9244 (CT5284)	39B3-4	BL167 BL3183	Df(2L)TW161 Df(2L)TW1	wt trachea tracheal phenotype
i261:bp1577 i205:bp2069	in frame	2596bp		CG11525 (CT36409)	100D1	BL3369	Df(3R)awd-KRB, e* ca1	general defects including trachea
i115: bp1393	in frame	4020bp		CG14792	30B9-2	BL10572	P{w^{+mC}=lacW}l(2)k05424^{k05424}	wt trachea
i12: bp1360	in frame	2315bp		Black Cells CG5779	55A1	BL1362	Bc^{1}	wt trachea
i17: bp1219	in frame	1857bp		CG12703	18E1-F5	EP(1)1116 EP(1)1344		wt trachea wt trachea
i25: bp508	in frame	1724bp		CG2944	41F3			
i238:bp 456 i90: bp759	in frame	1654bp		CG2713	3D6-C	BL733 BL734	Df(1)w258-42, y^{1} Df(1)w258-45, y^{2} sn^{3}	wt trachea wt trachea
i199: bp94 i180: bp161	i199 in frame i180 out of frame	1218bp		CG10189	37F4-5	BL3785	Df(2L)VA17, DdcVA17 pr^{1}	tracheal phenotype
i175:bp630	out of frame	3042bp		CG15740	10F			
i182: bp1227	out of frame	1518bp		CG14939	33B1			

i7	SEG:AE003433
i118	SEG:AE003512
i159	SEG:AE003492
i169	SEG:AE003512
i176	SEG:AE003633
i232	SEG:AE003809
i187	no hits
i226	

3. DISCUSSION

Vertebrate FGFRs activate the Ras MAPK pathway via the docking molecule FRS2 which is thought to recruit the Sos/Grb2 complex to the activated receptor (Wang et al. 1996; Kouhara et al. 1997). In *Drosophila*, a distant relative of FRS2 exists, but is not required for the transduction of the FGF signal (A. Michelson, personal communication). In contrast, the adaptor molecule Dof is required for this process. How Dof links FGFRs to the Ras MAPK pathway or to any other potential signaling pathways employed by the FGFRs is unknown. In this project, a yeast two-hybrid screen was conducted with the aim of identifying interacting partners that Dof uses to link FGFRs to their downstream signaling pathways. Amongst a large number of potential candidates that were isolated, three candidates provide particularly informative cues about Dof's role *in vivo*. Firstly, one of the FGFRs, Heartless, was isolated. Subsequent work showed that Breathless, the second FGFR in Drosophila, could also bind to Dof in the yeast two-hybrid system. Second, Dof itself was isolated, suggesting that homodimerization might play a role in Dof's *in vivo* function. Third, Ubc9, a SUMOlating enzyme, was identified as a binding partner of Dof, and the mutation of a putative SUMOlation site in Dof abolished its interaction with the receptors in the yeast two-hybrid system. Besides the receptor Heartless, a number of other molecules that have been suggested to be involved in signalling pathways were isolated. Other molecules, like a CDEP homologue may provide a direct link to the cytoskeleton, implicating Dof in the activation of pathways other that the MAPK pathway.

3.1 Summary of the yeast two-hybrid screen

1.3 x 10^6 transformants from a *Drosophila* cDNA expression library derived from 0 – 18 hour old embryos were used to screen for yeast two-hybrid interactions with the Dof deletion construct DofΔC. 269 clones activated the markers lacZ and His3. Plasmid DNA was isolated from these 269 positive colonies and retested for His3 activation with DB-DofΔC and for autoactivation. 128 clones did not autoactivate but still activated the His3 reporter. This group was called the group of "His3+ positives". The 5 prime end of each clone was sequenced and the clones were subsequently organized into 15 groups of transcripts that were represented more than once, and 34 single copies. The longest clone from each group and each of the single copies was subsequently tested for activation of the more stringent reporter, Ade2. Eight groups and three single clones still activated the Ade2 reporter, and these clones were called the group of "Ade2 positives".

The strain PJ694-a was used for the yeast two-hybrid screen because it minimizes false positives that have an affinity for a specific promoter sequence. Thus, in contrast to earlier yeast two-hybrid reporter strains, a different promoter sequence is placed downstream of each reporter, ensuring that if all three reporters are activated, this is not due to an affinity of the interactor with the promoter sequence. Unfortunately, and for unknown reasons, the Ade2 reporter has been shown to be particularly stringent and according to some reports, does not reflect weak yeast two-hybrid interactions (Phillip James, personal communication). Thus, while the differentiation between Ade2 and His3 positive groups is one way to distinguish between true and false interactors, the His3 positive clones were nevertheless considered to be potential binding partners of Dof.

Three Ade+ candidates and eleven His+ candidates isolated in the yeast two-hybrid screen have transcripts that are out of frame with the activation domain sequence. This would mean that a different reading frame from the wild-type open reading frame is read by the translational machinery. It has been observed, however, that during translation, a certain amount of "slippage" of the ribosome over stop codons or any other sequence can occur, resulting in a shift of the reading frame and the translation of a protein that is the same as that of the wild type sequence. The candidate group of transcripts coding for the homologue of *rps10* are all out of frame. In fact, both incorrect frames are represented, but no in-frame clone is represented in the group. Another candidate, clone i8, which corresponds to the transcript of the *sinuous* gene, is also out of frame. Fusion proteins in both other frames were subsequently tested, but neither of these interacted with Dof in the yeast-two hybrid system. However, one of the out-of-frame His+ clones isolated in this screen, Rack1, has been shown to immunoprecipitate with Dof (M. Affolter, personal communication). Taken together, if a transcript is out of frame with the activation domain sequence, this does not immediately mean that the interaction observed in the yeast two-hybrid system is false.

3.2 Interaction of Dof with FGFRs in the yeast two-hybrid system

Dof was found to interact with both FGFRs in the yeast two-hybrid system. In the screen itself, five transcripts corresponding to cytoplasmic domain of Heartless were isolated. The second *Drosophila* FGFR, Breathless, was subsequently also shown to interact with Dof in the yeast two-hybrid system. It is possible that Breathless was not isolated in the yeast two-hybrid screen because its transcripts are not highly represented in the expression library that was used or that the intron in the Breathless transcript was not efficiently spliced.

Regions in the FGFRs that contact Dof in the yeast two-hybrid system

The crystal structure of the cytoplasmic domain of human FGFR1 (Mohammadi et al. 1996) is depicted by a ribbon diagram in Figure 3.1A. Given the high conservation of the FGFRs (except for the juxtamembrane region, see below), it is likely that the cytoplasmic domains of Breathless and Heartless have a similar structure and the crystal structure of FGFR1 can therefore be used to illustrate which domains of Heartless and Breathless are contacting Dof in the yeast two-hybrid system.

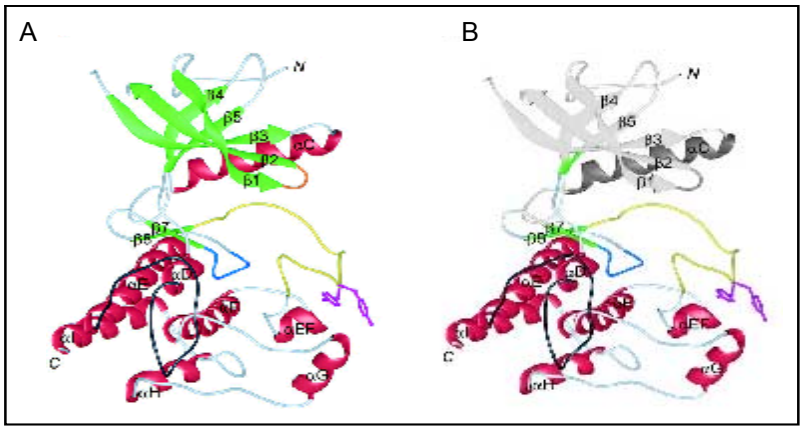


Figure 3.1 A and **B** are ribbon diagrams of the cytoplasmic FGFR1 structure, which does not include the juxtamembrane region and the last few residues of the C-terminus (adapted from Mohammadi et al. 1996). The α helices are shown in red, the β strands in green, the nucleotide-binding loop in orange, the catalytic loop in blue, the activation loop in yellow , the kinase insert in black, and the side chains of Tyr-653 and Tyr-654 in purple. In B, the region that is colored is equivalent to that encoded by i79, the shortest Heartless transcript isolated in the yeast two-hybrid screen.

The longest Heartless construct isolated in the screen corresponds to the complete cytoplasmic domain, including the juxtamembrane region, which is not shown in the figure because its crystal structure was not solved. The shortest Heartless construct (i79) and its Breathless equivalent that bind to Dof encode the C-terminal half of the cytoplasmic domain, (shown in color in Figure 3.1A) which contains most of the kinase domain, but not the ATP binding site (shown in orange in Figure 3.1A). This binding behaviour stands in contrast to FRS2, the adaptor molecule of vertebrate FGFRs, which binds to the juxtamembrane region of FGFR1 (Xu et al. 1998; Ong et al. 2000). Interestingly, the juxtamembrane region of Heartless and Breathless is not highly conserved with that of FGFR1.

Regions in Dof that contact the FGFRs in the yeast two-hybrid system

Requirement of the N-terminus of Dof for binding to the FGFRs

In order to find out which part of Dof is binding to the FGFR, the cytoplasmic domain of Heartless was co-expressed with a panel of Dof deletion constructs and assayed for activation of the His3 and Ade2 reporters. The whole C-terminal half of Dof (construct 405) could be removed without affecting the interaction. In contrast, when parts of the N-terminus were removed from Dof, interaction did not occur. One construct behaved in an unpredicted way: Construct 427 contains deletion at its N- and C-terminus, but still interacts with the receptor. This might reflect the fact that cis interactions amongst regions of the Dof molecule are regulating its binding affinity to interacting partners.

Behaviour of the C-terminus of Dof

The screen was conducted with a Dof construct containing a C-terminal deletion (DofΔC), but a subsequent test showed that full length Dof also binds to the entire

cytoplasmic domain of Heartless (Figure 2.7). However, this binding is not strong enough to activate the stringent reporter Ade2, indicating that the C-terminus of Dof might obstruct its contact with the receptor. In fact, the minimal receptor domains of both Heartless and Breathless could only interact with Dof on Ade2 when the whole C-terminal half was deleted (construct 405, Figure 2.7). If these yeast two-hybrid interactions reflect the *in vivo* binding of Dof to the receptor, they indicate that the N- and C- terminus of Dof can modulate its binding affinity to the receptor.

An amino acid substitution that abolishes interaction between Dof and the receptor

In a separate set of experiments, the third putative SUMOlation site in Dof was mutated by substituting the amino acid lysine at position 297 for an Arginine. This mutation was introduced into the construct DofΔC and resulted in a complete disruption of interaction between the two FGFRs and Dof (Figure 2.17). Since this amino acid substitution did not affect the binding of Dof to any other interaction partner it is likely that no major structural changes have been induced by the K297R mutation. If indeed this amino acid substitution has created a molecule that cannot be SUMOlated it could mean that SUMOlation is required for Dof's *in vivo* function. The requirement of a SUMOlation site for protein-protein interaction has been previously observed in the SUMO target RanGAP1. RanGAP1, the GTPase-activating protein for the Ras-related GTPase Ran that is involved in nucleocytoplasmic transport (Mahajan et al. 1997; Matunis et al. 1996) is a target for SUMO1 modification. RanGAP1 forms a stable complex with RanVP2/Nup358 at the nuclear pore complex. The association of RanGAP1 with RanBP2 depends on the modification of RanGAP1 with a single SUMO1 molecule that is attached to lysine 526 in RanGAP1 (Matunis et al. 1998). SUMO itself does not bind to RanBP2 and domain mapping of the binding site of RanBP2 in RanGAP1 has led to the speculation that SUMOlation either contributes to the RanBP2 binding or alters the conformation of RanGAP such that a hidden binding site becomes accessible.

Thus, Dof might require SUMOlation for its binding to FGFRs in a similar fashion in which RanGAP requires it to bind to RanBP2. *Drosophila* embryos that are homozygous mutant for Ubc9, the SUMOlating enzyme, develop normal tracheae and mesoderm (see discussion of Ubc9), however, it is possible that a strong maternal component of Ubc9 masks the requirement of Dof to be SUMOlated. It therefore remains to be determined whether SUMOlation of Dof is required for its *in vivo* function.

Dof might only interact with a non-phosphorylated receptor

The interaction between Dof and Heartless that led to its isolation in the yeast two-hybrid screen occurred when Heartless was expressed as fusion protein with the Gal4 activation domain. However, when the full length cytoplasmic domain is expressed as a Gal4 binding domain fusion, interaction with Dof no longer occurs.

Although it is commonly observed that an interaction in the yeast two-hybrid system occurs in one direction, but not the other (see, for example, Dof dimerization experiments), the lack of binding of Dof to Heartless when it is fused to the binding

domain may possibly be attributed to the receptor's ability to autophosphorylate itself. Xu H. et al. (1998) showed that when FGFR-1 cytoplasmic domains are brought together at the Gal4 binding sites, autophosphorylation occurs. It is therefore plausible that Dof does not bind to DB-Heartless because it is phosphorylated. Thus Dof might only be able to bind to activation domain fusions of Heartless and Breathless because they are not phosphorylated. Indeed, when the ATP binding domain of a DNA binding fusion of Breathless is removed, a weak activation on His3 can be observed (Figure 2.5), indicating that the reason why the full length cytoplasmic domain of Heartless does not interact with Dof in this direction might be because it is phosphorylated.

While further experiments, such as co-immunoprecipitations, have to confirm that Dof does bind directly to the FGFRs, the yeast two-hybrid data imply that Dof does interact directly with the receptor, most likely with the unphosphorylated receptor. When Dof is expressed in Schneider cells in the presence of a consitutively active Breathless receptor, it does not co-immunoprecipitate (R.Wilson, personal communication), supporting the idea that if Dof does bind to the FGFRs, then it only binds to the un-phosphorylated receptor. This observation has implications for the function of Dof and could mean that Dof binds to the receptor when it is inactive and upon activation, is released from the receptor and can thereby activate downstream targets.

Thus, the mechanism by which it transduces the FGF signal is different from docking molecules like FRS2, which bind to the phosphorylated receptor and link it to Sos/Grb2 and subsequently to the Ras MAPK pathway. It is also possible that Dof provides structural support for the inactive receptor, in the manner of a co-receptor. Or it is even possible that Dof binds to receptor because it plays a role in a process unrelated to the receptor's signaling function, like in its transport to the plasma membrane.

3.3 Dof self-association

Prior to conducting the yeast two-hybrid screen for proteins that interact with Dof, co-expression experiments of DB-DofΔC with AD- DofΔC had demonstrated that Dof can homodimerize. In the yeast two-hybrid screen using DB-DofΔC as bait, Dof was isolated as a candidate. In order to test the *in vivo* relevance of this dimerization behaviour, the yeast two-hybrid system was employed to define the regions of the molecule that lead to dimerization.

Regions of the Dof molecule that mediate dimerization

Combinations of deletion constructs were co-expressed and assayed for interactions with each other. This data was used to narrow down the interacting domains to the two separate parts of the molecule, that were vicariously defined as m#1 and Hd#2. Briefly, and as discussed in detail in Chapter 2.4, region m#1 is required for the N-terminus to bind to itself, but is also required for binding to the C-terminal domain, Hd#2. On the other hand, Hd#2 cannot bind to itself. Finally, in some co-expressions, it appeared that the C-terminus weakens the self-association.

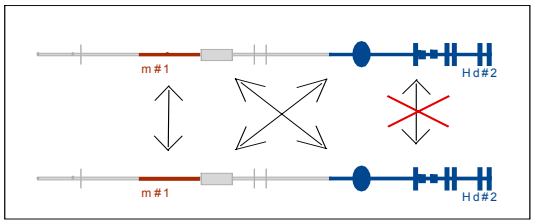


Figure 3.2 The arrows in this figure show the domains of Dof do or don't bind to each other in the yeast two-hybrid system

Since the third potential SUMOlation site lies within the region m#1, a mutation in the SUMO acceptor site, namely the lysine at position 297, was introduced. The reasoning behind this was that yeast SUMO dimerizes in the yeast two-hybrid system (I. Schwienhorst, personal communication). It would therefore be possible, given that Ubc9 binds to Dof, that the SUMO moiety was dimerizing, thus causing the Dof molecule to dimerize. However, when the mutation is introduced into the BD-DofΔC and AD-DofΔC constructs, they still interact, indicating that the SUMOlation site does not mediate interaction between the two m#1 domains.

The matrix of Dof deletion interactions narrows down the domains of Dof that are required for an interaction to occur, but it is still necessary to test whether domain m#1 expressed on its own can bind to itself and to Hd#2. However, if interaction does not occur in the yeast two-hybrid system, it would remain possible that a structural element is missing when m#1 is expressed on its own.

Although dimerization of the Dof molecule is observed in the yeast two-hybrid system, it remains to be determined whether this behaviour is important for its function *in vivo*. An interallelic complementation test that confirms its relevance has already been done by R. Wilson (personal communication). The construct 430 does not rescue tracheal development and neither does the construct 405. Construct 430 lacks the m#1 region, whereas construct 405 lacks the C-terminus. These two constructs dimerize in the yeast two hybrid system, and when expressed together in the trachea, can rescue the *dof* mutant phenotype. Further experiments will have to confirm that rescue is occuring via dimerization, and not via intermediate molecules, but this experiment indicates that homodimerization might be important for the *in vivo* function of Dof.

Possible conformations of the dimer

In the yeast two-hybrid system, dimerization of the Dof molecule that leads to reporter activation must occur in trans for the activation and binding domains to be brought together. Thus, *in vivo*, one Dof molecule might bind in trans to a second Dof molecule via the domains m#1 and Hd#2. The conformation of two Dof molecules only binding in trans are shown in Figure 3.3 A-C. It is also possible for any of these combinations to form a multimer consisting of Dof dimers. However, it is equally likely that the molecule is folded on itself, resulting from the C-terminus binding to the N-terminus. This could occur at the level of a single molecule (Figure 3.3 D), or several molecules could be

folded upon themselves and at the same time bind to a second Dof molecule (Figure 3.3 E and F).

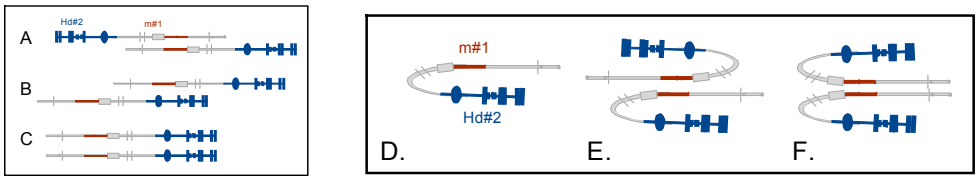


Figure 3.3 This figure show models of possible homodimers that can result from Dof self-association.

Cis interactions

As discussed above, it is possible that the C-terminal domain Hd#2 binds in cis to the N-terminal domain m#1. This possibility is strengthened by the fact that in yeast two-hybrid assays, the full length Dof molecule often interacts less strongly with binding partners than the C-terminal deletion construct, meaning that the C-terminus might block an interaction domain (see Dof-FGFR interactions or Dof dimerizations). There is no proof, however, that this is functionally relevant, especially since C-terminal deletions of the Dof molecule, when expressed in the trachea, do not give a constitutively active phenotype (for example, excessive branching, as seen in the case of constitutively activated Breathless expressed in the trachea).

Dimerization as a mechanism in signal transduction

Dof might assemble a multicomponent signaling complex by dimerizing. It also could use self-interaction in cis or trans to regulate its activity. Self-interaction has been observed in various molecules, including the scaffold protein Ste 5 in *Saccharomyces cerevisiae*. Its dimerizing behaviour has been characterized and can serve as an example of how self-interaction can mediate signaling. Dimerization has been shown to be essential for Ste5's function (reviewed in Garrington and Johnson, 1999), which is to act as a scaffold for the MAPK cascade components in a yeast pheromone response pathway. An N-terminal Ring-H2 domain (Inouye et al. 1997) mediates Ste5 dimerization. This domain also binds to Ste4, the Gβ subunit of the heterotrimeric guanine nucleotide-binding protein that is released from the pheromone receptor upon its activation. By artificially inducing dimerization in mutant and wild type Ste5 constructs, Inouye et al. showed that a complementation of *ste4*Δ*ste5*Δ occurs if the molecule lacking a functional Ring-H2 domain can dimerize artificially (by fusing the molecules to GST, which dimerizes). However, complementation of *ste4*Δ*ste5*Δ does not occur when the GST construct contains a functional Ring-H2 domain, indicating that the Ring-H2 domain might inhibit Ste5 from dimerizing, but upon Gβ binding, promote a conformational change that permits Ste5 dimerization.

Dof does not have a Ring-H2 domain and has not been shown to bind to components of the MAPK pathway, but the mechanism of Ste5 scaffold function via dimerization provides a basis for understanding the possible mode by which Dof transduces a signal from the FGF receptor. It is possible that in a similar manner to Ste5, Dof uses dimerization as a means to assemble a multicomponent signaling complex to

channel specific sets of signaling molecules into one direction. Like the control of Ste5 by the Gβ subunit, dimerization of Dof could be controlled by other molecules which induce it to self-associate upon receptor activation. Dof is not evolutionarily conserved and it is interesting to note that Ste5 is not evolutionarily conserved either, giving rise to the idea that signaling networks of evolutionarily highly conserved molecules might be mixed and matched by species specific intermediate proteins to give rise to the unique phenotype of each species.

3.4 Potential role of Dof SUMOlation

The SUMOlating enzyme Ubc9 was isolated in the yeast two-hybrid screen as a binding partner of Dof. It has been shown that SUMOlation can mediate the intracellular localization of proteins, as well as their interaction with other molecules (reviewed in Melchior 2000). Experiments from this work suggest that Dof might be SUMOlated. A mutation in Dof's third putative SUMOlation site abolishes its yeast two-hybrid interaction with the FGFRs. The third putative SUMOlation site lies in the region that is deleted in construct 430. The region deleted from this construct is particularly relevant for the *in vivo* function of Dof because this construct does not rescue the *dof* mutant phenotype. In fact, the protein encoded by construct 430 is predominantly found in the nucleus (R. Wilson, personal communication) which contrasts with the cytoplasmic localization of wild-type Dof (Vincent et al. 1998). Since SUMOlation has been shown to mediate protein localization, it is possible that this site mediates Dof localization to the plasma membrane or its retention in the cytoplasm. The relevance of this mutation remains to be tested *in vivo*, namely by assaying whether the SUMO site mutant can rescue the *dof* mutant phenotype. Furthermore, it remains to be established whether Dof is SUMOlated *in vivo*. Ubc9 still interacts with the SUMOlation site mutant in the yeast two hybrid assay. This has been observed for other Ubc9 targets, like RanBP2 (Lee et al. 1998), which still coprecipitates with Ubc9 when it is not SUMOlated. Tracheal and mesoderm development in Ubc9 mutant embryos was found to be as in wild type embryos. However, this does not mean that Ubc9 SUMOlation is not required, since the possibility remains that the mutant embryo inherits a maternal supply of Ubc9.

So far, the majority of SUMO targets that have been described function in the nucleus and have to be SUMOlated to get there. If Dof does indeed require SUMOlation for its function, it could mean two things, namely that it might have an unknown function in the nucleus or that SUMOlation might regulate its cytoplasmic localization, retention or interaction with other molecules.

3.5 Conclusion

In the yeast two-hybrid screen for interacting partners of Dof, a number of candidates that give clues to how Dof is functioning *in vivo* were identified. Many of these candidates, especially those that have been implicated in signaling functions or those which might play a role in signaling to the cytoskeleton, remain to be characterized. The three candidates studied in more detail in this project raise some interesting possibilities for Dof's function in FGF signal transduction. Dof's binding behaviour with the FGFRs in

the yeast-two hybrid system indicates that it might bind to the non-phosphorylated receptor and that regions within the molecule affect the affinity with which it interacts. The binding of Dof to the FGFRs might be regulated *in vivo* by its post-translational modification with a SUMO moiety at its N-terminus, since the putative SUMOlation site in Dof is required for its interaction with the FGFRs in the yeast two-hybrid system. Finally, since Dof self-associates in the yeast two-hybrid system, it is possible that dimerization is important for its function.

4. MATERIALS AND METHODS

4.1 Materials

Saccharomyces cerevisiae (S.cerevisiae)

The yeast strain PJ69-4A (gift from Phillip James) was used for all yeast two-hybrid assays. The genotype of PJ69-4A is *MATa trp1-90 leu2-3,112 ura3-52 his3-200 gal4△ gal80△ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ.*

Escherichia coli (E. coli)

JM107 (endA1, gyrA96, thi, hsdR17, supE44, relA1, △(lac-proAB),(F', traD36, $proAB^+$, $lacI^q$ lacZ△M15) (Yannisch-Perron et al. 1985)

DH5α (supE44 Δ(lac ZYA-argF)U169 (Φ80 lacZDM15) hsdR17($r_k^-m_k^+$) recA1 endA1 gyrA96(NaI^r) thi1 relA1)

Drosophila melanogaster

Wild type flies were w^{1118}. Unless mentioned otherwise, strains were obtained from Bloomington, Indiana. A description of balancers and markers can be found in Lindsley and Zimm (1992).

DNA constructs

Constructs used in this work are described in Appendix IV.

Antibodies

The following primary antibodies were used: monoclonal mouse mAB2A12 (provided by N. Patel) was used at 1: 10, mouse anti-FasciclinIII (provided by R. Smith) was used at 1:10, and rabbit anti-Eve (provided by M. Frasch) was used at 1:2000 .

Oligonucleotides

Oligonucleotides were purchased from Eurogentech. The lyophilized pellet was resuspended in water to a concentration of 100pmol/μl and stored at –20°C.

Computer Software and digital Photography

Digital pictures were taken of images under Axiophot Photomikroscop (Zeiss) with the ProgRes 3008 (Kontron Elektronik) camera. Pictures were edited using Adobe Photoshop (Adobe systems) software. Figures were drawn using Canvas 7.0 (Deneba

Systems) software. DNA sequence alignments and analysis were made using VectorNTI (InforMax Inc.) software.

Chemicals and Reagents

Tris was purchased from Applichem, Aquaphenol purchase from Appligene. Roche Diagnostics GmbH supplied herring sperm DNA (used for hybridization buffer), Nitroblue tetrazolium (NBT), 5-bromo-4-cholor-indoxylphophate (BCIP), unlabeled nucleotides, Hexanucleotide Mix, DIG-DNA labelling mix, anti-DIG antibody, Proteinase K and Expand Enzyme Mix. Agarose Electrophoresis Grade, Casein Hydrolysate, Yeast Extract, and Select Agar were purchased from Gibco BRL. Acelylated Bovine Serum Albumin (BSA) was purchase from New England Biolabs. Heparin was bought from Serva and Sigma supplied Tween20, tRNA, RNase. Ampicillin, and Tetracycline. Amino Acids were purchase from Sigma. Molecular Weight Markers. The λ/HindIII Marker and the 1kb ladder were purchased from NEB. Restriction enzymes and their reaction buffers were supplied by New England Biolabs and Roche Diagnostics GmbH. Unless otherwise mentioned, all other chemicals were purchased from Merck (Darmstadt), Sigma (Deisenhofen) or Roth (Karlsruhe).

4.2 Methods

Sequence similarity and motif searches

Sequence similarity searches were conducted using the Basic Local Alignment Search Tool (BLAST) developed at NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/). Searches for annotated Drosophila genomic sequences, Drosophila ESTs and Drosophila deficiency and P-element insertion strains were conducted using FlyBase (FlyBase (1999);http://flybase.bio.indiana.edu/) Motif searches were conducted using PFAM at the website http://www.sanger.ac.uk/Software/Pfam/search.shtml (Bateman et al. 1999).

Embryo collections and maintenance of flies

The flies (Drosophila melanogaster) were maintained under standard conditions (Ashburner, 1989; Wieschaus and Nüsslein-Volhard, 1986) and the embryos were collected under standard conditions (Ashburner, 1989; Wieschaus and Nuesslein Volhard, 1986).

Molecular biology

The methods used for cloning were those descirbed in Sambrook et al. (1989)

DNA sequencing

DNA was sequenced by the primer extension dideoxynucleotide extension method (Sanger, 1977) using a Dye-Terminator Kit (Perkin-Elmer) and the sequencing reaction

was read by a 373A ABI-Sequencer (Applied Biosystems). DNA for the sequencing reaction was prepared either by the alkaline miniprep procedure or by using the Jetstar Plasmid Midiprep Kit. DNA sequences were aligned using the computer software Vector NTI (InforMax Inc.).

Competent *E. Coli* cells

Electrocompetent cells were prepared using the method developed by Hanahan et al. (1995). Chemical competent cells were purchased from GibcoBRL.

In situ hybridization to embryos with DIG-labeled RNA probes

In situ hybridizations were conducted as described by Tautz and Pfeifle (1989) Since DIG-labeled RNA probes were used, hybridization was carried out at 55°C.

Digoxigenin (DIG) labeling of RNA

DIG labeled RNA probes were prepared by using a PCR product as a template for the transcription reaction. The DNA template was made using primers containing the core sequence of T7, T3 or SP6 RNA polymerase binding sites. For antisense probes of inserts in the vector pACT2, the primers #130 and #131 were used. For sense probes of inserts in the vector pACT2 (used as a control), the primers #132 and #133 were used. For antisense probes of inserts in the vector pOT2, primers #165 and #170 or #166 and #170 were used.

primer#	name	vector	primer sequence (5' to 3' direction)	Polymerase binding site
#130	act25'	pACT2	tggtcatatggccatggagg	none
#131	T7αsense	pACT2	taatacgactcactatagggaggacgattcatagatctctc	T7 polymerase
#132	act23'	pACT2	tgcacgatgcacagttgaag	none
#133	T7sense	pACT2	taatacgactcactatagggaggtcatatggccatggaggc	T7 polymerase
#165	pOT2 p3	pOT2	catacgatttaggtgacactatagcgttagaacgcggctac	SP6 polymerase
#166	pOT2 p4	pOT2	attaaccctcactaaagggacgttagaacgcggctacaat	T3 polymerase
#170	T7	pOT2	taatacgactcactatagggaga	none

Table 4.1: This table lists the primer used to generate PCR products used as templates for making DIG-labeled probes.

The PCR reaction was assembled in a total volume of 25µl using the following components: 20ng plasmid DNA, 25pmol each primer, 1µl 10mM dNTP, Expand PCR Buffer (2.5mM $MgCl_2$) and 1µl Expand Taq polymerase (3.5 units) added last. The PCR conditions were 2 minutes at 94°C and then 30 cycles of 94°C (10 seconds), 68°C (30 seconds) and 72°C (2 minutes). The concentration of PCR product was estimated on an agarose gel using the λ/HindIII marker as a reference. For the DIG labeling reaction in a final volume of 20µl, 100-200ng PCR product was added to 2µl DIG RNA Labeling Mix 10X (10mM ATP, 10mM CTP, 10mM GTP, 6,5mM UTP, 3.5mM DIG-11-UTP, Roche Diagnostics GmbH), 2µl 10X Transcription Buffer (400mM Tris-HCl, pH 8.0 (20°C), 60mM $MgCL_2$, 100mM DTT, 20mM Spermidine; Roche Diagnostics GmbH) 0.5µl

RNase Inhibitor (40U/µl, Promega), 2µl 100mMDTT and deionized water to a final volume of 20µl. After all reaction components had been assembled, 2µl RNA polymerase (40 Units) were added and the reaction was incubated at 37°C for 2 hours. Afterwards, the reaction mix was brought to 100µl with hybridization mix (1:5 dilution) and stored at –20. The probe was used for in situ hybridizations at dilutions ranging from 1:20 to 1:20 000, depending on the probe.

Immunohistochemistry on Drosophila embryos

For antibody stainings, the method described by Mitchison and Sedat (1983) was used. PBS (100mM Na-Phosphate pH7.1, 130mM NaCl) was used instead of Buffer AM and 10%BSA was used instead of 10%NGS. The biotinylated secondary antibodies were visualized with Vectastain-ABC-Kit (purchased from Camon-Labor-Service), using the manufacturer's specifications

Embedding embryos in Araldite

For microscopy, embryos were embedded in Araldite. Embryos in PBS (100mM Na-Phosphate pH7.1, 130mM NaCl) were rehydrated by short 30%, 50%,90% and 96% EtOH washes. This is followed by a 10 minute incubation in 100% water-free EtOH (EtOH stored with Molecular Sieves, Sigma M-2010, 1/8 inch pellets) and then by a gravity wash in water-free Acetone. About 100µl Araldite:Acetone 1:1 mix (reagents for Araldite supplied by Sigma) was then added to the embryos and these were then transfered into a shallow plastic lid containing pure Araldite.

Isolation of DNA from adult flies

20 flies were homogenized in 250µl Buffer A (0.1M Tris pH9, 0.1M EDTA, 1%SDS). Another 250µl Buffer A was subsequently added and the solution was incubated at 70°C for 30 minutes followed by 5 minutes at 0°C. Then 70µl 8M KoAc was added and the solution vortexed. After vortexing, the solution was incubated for 30 minutes at 0°C and then centrifuged in a 12000 bench top centrifuge for 10 minutes. The supernatant was transfered to a fresh 1.5ml Eppendorf tube and the genomic DNA was precipitated with 0.7 volumes Isopropanol. The pellet was resuspended in 200ml TE and the suspension was extracted twice with Phenol/Chloroform/Isoamylalcohol 25:24:1. The genomic DNA was then precipitated with 2.5 volumes Ethanol and 1/10 volume 3MNaOAc pH 5.2. The pellet was washed with 70%EtOH and brought to an approximate concentration to 500ng/µl.

PCR

The PCR was carried out in a UNO Thermoblock (Biometra). The reaction components were: 20pmol each primer, 20ng template DNA (for analytical purposes) or 1µg template DNA if the template was used for cloning, 1xPCR buffer (supplier), 10pmol dNTPs and 0.75µl Expand enzyme mix (Expand High Fidelity PCR System, Roche) in 50µl reaction

volume. The reaction conditions were optimized for each primer pair, but essentially, underwent 30 cycles of 1 minute at 95°C, 1 minute at the annealing temperature optimized for the specific primer pair, and 1 minute (or more, depending on size of PCR product) extension at 72°C.

4.3 Methods for yeast two-hybrid assays

Media Recipes

Most Recipes and Protocols used for experiments with *Saccharomyces cerevisiae* were adapted from protocols provided by the Gietz home page http://www.umanitoba.ca/faculties/medicine/units/biochem/gietz/.

Liquid YPAD (Yeast Extract-Peptone-Dextrose plus Adenine medium): 6g yeast extract, 12g Peptone, 12g Glucose, 60mg adenine hemisulphate, 600ml distilled water. Sterilise by autoclaving 20 minutes at 15 lb/sq.in. on liquid cycle. For solid medium: prepare liquid YPAD and add 10g Bacto-agar (Difco) before autoclaving.
SC drop-out Medium (Synthetic complete Drop Out Medium): 4.0g Difco Yeast Nitrogen Base (without amino acids), 12g Glucose, 0.4g Synthetic Complete Drop Out Medium Mix (see below), 600 ml distilled water. Mix all ingredients in water and adjust the pH to 5.6 with 10N NaOH. Sterilise by autoclaving 20 minutes at 15 lb/sq.in. on liquid cycle. For solid medium, add 10g Difco Bacto Agar before autoclaving.
Synthetic Complete drop-out Medium Mix: 2.0g Adenine Hemisulfate, 2.0g Arginine HCl, 2.0g Histidine HCl, 2.0g Isoleucine, 2.0g Lysine HCl, 2.0g Methionine, 2.0g Tyrosine, 2.0g Threonine, 2.0g Serine, 4.0g Leucine, 3.0 g Phenlyalanine, 3.0 g Tryptophane, 1.2g Uracil, 9.0g Valine. Store mix at room temperature. Note: this mix is made by omitting the component for which is to be selected. Thus, "SC – Leu" means that Leucine has been omitted

Frozen Cell transformation method

For standard transformations of *S. cerevisiae*, a variation of the frozen yeast method was used (Dohmen et. al 1991). To the frozen cell suspension (see below) 0.1-5 µg plasmid DNA and 50µg herring sperm DNA (5µl of 10mg/ml Stock, provided by Clontech) in a maximum volume of 20µl were added. This suspension was mixed for 2 minutes or until cells were thawed. 1ml Solution 2 (40% PEG1000 (Roth#0150.1), 0.2M Bicine-NaoH pH 8.35 (Sigma B8660);sterilized by autoclaving 20 minutes at 15 lb/sq.in. on liquid cycle) was added and mixed by inverting the tube 2-3 times. This suspension was incubated at 30°C for 1hour. The cells were then collected by spinning them down at 3000g for 5 seconds and resuspending the pellet in 200µl Solution 3 (0.15M NaCl, 10mM Bicine-NaOH pH8.35 (Sigma B8660);sterilized by autoclaving 20 minutes at 15 lb/sq.in. on liquid cycle) and plating the suspension onto SC omission medium. Transformants could be seen after an incubation at 30°C for 2 days.

Preparation of cells for the frozen yeast method

Cells (10ml per desired transformation) were grown overnight at 30°C in YPAD to an OD_{600} of 0.6-0.8 (which amounts to about 0.6-1 x 10^7 cells/ml). Cells were washed in 0.5 volumes Solution 1 (1.0M Sorbitol, 10mM Bicine-NaOH (pH 8.35) Sigma B8660, 3% Ethylene glycol, 5% DMSO (Sigma D8418); sterilized by autoclaving 20 minutes at 15 lb/sq.in. on liquid cycle) and resuspend in 0.02 volume Solution 1. 200µl aliquots were frozen slowly by placing cells at –70°C within a styropore box.

Description of pACT2 Drosophila embryonic expression library

The cDNA library in pACT2 (Durfee et al. 1993) used for the yeast two hybrid screen was a gift from S. Elledge. The cDNA for the library was made from mRNA from 0-18 hour old embryos and was size selected to be larger than 1kb. The library was received in the lamda phage λACTand was reported to have between 1x10^7 and 1x10^8 total recombinants, amplified only once from the packaging (S.Elledge, personal communication) The titer was 2x10^9 PFU/ml (the bacterial strain used for this measurement was DH5α). The strain JM107/λKC was used to convert the phage vector into plasmid vector (see below). Information about the priming method for the constuction of the library was not given. The sequence of the clones isolated in the screen have the following 5 prime and 3 prime vector ends: "ggggatccgaattcggcacgagxxxxx.....xxxxxttttctcgag" where "x" represents the sequence from the cDNA insert.

Automatic subcloning conversion of λACT into pACT

The Drosophila cDNA library in λACT (see above) was converted from a phage library into the pACT plasmid library by infecting the bacterial strain JM107/λKC, a kan^r lamda lyosgen containing the cre gen, with the λACT phage library. Upon infection of JM107/λKC with the library, the plasmid is excised via cre-mediated lox recombination, thus creating Amp^r colonies. The detailed protocol (adapted from Durfee et al.) is described below.
A single JM107/λKC colony was used to inoculate 50ml LB-kan (10µg/ml) 0.2% Maltose. The culture was grown overnight at 37°C on a shaker at 200 cycles/minute. The next day, the pellet of cells from 2ml of overnight culture were resuspended in 2ml 0.01M $MgCl_2$ and subsequently infected with 300µl λACT phage solution (equivalent to about 5x10^8 PFU according to titer determined with DH5α, see above) and left to incubate for 30 minutes at 30°C without shaking. 2ml LB was then added and the culture was shaken for 1hr at 30°C. The total volume of 4.05ml was plated onto 10 X 150mm LB-amp 50µg/µl plates and grown at 37°C overnight. The total number of transformants was 1.873 x 10^{10}. This indicates that the titer determined using DH5α cells was an underestimation. The next day, plasmid DNA was prepared directly from the lysogens scraped from the 10 plates by growing them at 37°C to an OD of 1.1-1.2 in 6L terrific broth (4x1.5L) and then spinning them 20 minutes 8000rpm (Sorvall superspeed centrifuge - Du Pont). The pellet was vortexed and resuspended in 60ml Solution I (4x15ml). After vortexing, Solution II was added. The solutions were mixed by

vortexing and then Solution III was added. The solution was transfered to 50ml Falcon tubes and then spun in an Hereaus 20 minutes 4000rpm. The plasmid DNA was precipitated out of the supernatant at room temperature with 0.7 volume Isopropanol and recovered by spinning for 20 minutes at 4000rpm (Hereaus). The pellet was resuspended in 2ml H_2O and 2ml 5MLiCl. This was spun once again 20 minutes 4000rpm in an Hereaus. The pellet was resuspended in 10ml H_2O and 2.5-3volumes EtOH were added. The solution was spun for 20 minutes at 4000rpm and the pellet was resuspended in 3ml H_2O. 180µlEtBr (10µg/µl) was then, followed by 2.1ml 7.5M NH_3OAc, and then mixed. Then 4.8ml Phenol/Chloroform pH 8.0. was added, the solution vortexed and then centrifuged for 5 minutes. DNA was precipitated from the aqueous phase using 2.5 volumes EtOH and washing with 70% EtOH. The pellet was resuspended in 1.6ml water. The total yield of DNA was 16.8mg from $18x10^9$ transformants, which was half as much as the 2mg DNA yield from $1x10^9$ transformants reported by Steve Elledge (personal communication). 10 clones were picked off dilution plate (the plate used for calculating total transformants) and DNA was isolated from these clones and digested with HindIII digest. 3/10 clones were empty. 1/10 was not pACT2.

Transformation of cDNA expression library into AR2.9 and select of his+lacZ+ clones

The method used for transforming the pACT2 expression library into AR2.9 was adopted from Gietz and Schiestl (1995) and is summarized on the world wide web at http://www.umanitoba.ca/faculties/medicine/units/biochem/gietz/

The transformation of the library was conducted by scaling up the reaction components by 180X. 115.2ml overnight YPAD AR2.9 culture (4.44×10^9 cells) were pelleted and resuspended in 900ml YPAD. The suspension was grown for 7 hours at 30°C to an $O.D._{600}$ of 1.45. The suspension was then centrifuged and the pellet washed with an equal volume of millipore water. This suspension was centrifuged (3200rpm 5 minutes in Heraeus Sepatech centrifuge) and the millipore water was removed from the resulting pellet. This was followed by brief spin to get rid of residual water. The following transformation components (equal to a 180X scale-up) were added to the pellet: 43.2ml PEG50% (w/v), 6.48ml 1M LiAC, 900µl herring sperm DNA 10mg/ml (Clontech), 180µg pACT library DNA (in 600µl) and 12ml Millipore water. The suspension was incubated at 30°C for 30 minutes followed by 42°C for 35 minutes (shaking every 5 minutes). The pellet was isolated by centrifugation and resuspended in 67.5ml SC-ULH liquid medium and plated onto 70 SC-ULH selection plates (145mm). The total number of transformants (measured by plating a dilution series on SC-UL selection plates) was 1272150. The transformation efficiency was about 7×10^3 transformants per µg library DNA. The transformants were grown on SC-ULH plates for 7 days at 30°C and then replica plated onto SC-ULH plates containing 40µg/ml X-Gal. After incubation at 30°C for two days, 269 colonies had turned blue (that is, theywere expressing the lacZ reporter gene) and were streak purified onto SC-ULH plates. Plasmid DNA was isolated from the 269 yeast colonies, transformed into electrocompetent DH5α *E. Coli* cells and plated onto standard LB-Amp plates to isolate the insert in pACT (which confers Ampicillin resistance to the cells). DNA was prepared and digested with RI-XhoI to determine the

size of the plasmid. The plasmid was transformed back into the strain AR2.9 (to confirm the interaction with DofΔC) and in parallel into PJ69-4A/pGBDU (to test for autoactivation) and assayed for His3 activation on SC-ULH plates. All autoactivators were discarded. Sequence data of the 5prime end of all clones that grew on His3 was obtained using the primer with the sequence 5'ACCACTACAATGGATGATG 3'.

The strain AR2.9 used for the yeast two-hybrid screen

The construct "int.DB-Dof△C" was linearized with PvuI and transformed into PJ69-4A and transformants were selected for growth on SC-U plates. 10 clones were streak purified on SC-U. The expression of DofΔC from the integrated construct was tested by assaying for His3 activation upon co-expression with AD-DofΔC. Several of the 10 integration strains activated His3 upon co-expression with AD-DofΔC. The strain with which the best transformation efficiency could be achieved was selected and called AR2.9.

5. BIBLIOGRAPHY

(1999). "The FlyBase database of the Drosophila Genome Projects and community literature. The FlyBase Consortium." Nucleic Acids Res **27**(1): 85-8.

Adams, M. D., S. E. Celniker, et al. (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-95.

Alessi, D. R., M. Deak, et al. (1997). "3-Phosphoinositide-dependent protein kinase-1 (PDK1): structural and functional homology with the Drosophila DSTPK61 kinase." Curr Biol **7**(10): 776-89.

Allard, J. D., H. C. Chang, et al. (1996). "The SH2-containing tyrosine phosphatase corkscrew is required during signaling by sevenless, Ras1 and Raf." Development **122**(4): 1137-46.

Ashburner, M. (1989). *Drosophila - A laboratory handbook*. Cold Spring Haror Laboratory Press.

Ayer, S., N. Walker, et al. (1993). "Activation and repression of Drosophila alcohol dehydrogenase distal transcription by two steroid hormone receptor superfamily members binding to a common response element." Nucleic Acids Res **21**(7): 1619-27.

Bar-Sagi, D. and A. Hall (2000). "Ras and Rho GTPases: a family reunion." Cell **103**(2): 227-38.

Bateman, A., E. Birney, et al. (1999). "Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins." Nucleic Acids Res **27**(1): 260-2.

Bhat, M. A., S. Izaddoost, et al. (1999). "Discs Lost, a novel multi-PDZ domain protein, establishes and maintains epithelial polarity." Cell **96**(6): 833-45.

Blenis, J. (1993). "Signal transduction via the MAP kinases: proceed at your own RSK." Proc Natl Acad Sci U S A **90**(13): 5889-92.

Bate, M. (1993). The mesoderm and its derivatives. In *The Development of Drosophila melanogaster.* (ed. Bate, M., Martinez-Arias, A.) Cold Spring Harbor Lab., Plainview, NY; pp1013-1090

Brent, R. and M. Ptashne (1985). "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor." Cell **43**(3 Pt 2): 729-36.

Bray, D., 2000 *Cell Movements*. Garland Publishing, Inc., New York.

Campos-Ortega, J.A.; Hartenstein, V. (1997). *The embryonic development of Drosophila melanogaster*. Second Edition., Springer, Berlin

Chishti, A. H., A. C. Kim, et al. (1998). "The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane." Trends Biochem Sci **23**(8): 281-2.

Darnell, J. E., Jr. (1996). "The JAK-STAT pathway: summary of initial studies and recent advances." Recent Prog Horm Res **51**: 391-403.

Davis, S., P. Vanhoutte, et al. (2000). "The MAPK/ERK cascade targets both Elk-1 and cAMP response element- binding protein to control long-term potentiation-dependent gene expression in the dentate gyrus in vivo." J Neurosci **20**(12): 4563-72.

Dohmen, R. J., A. W. Strasser, et al. (1991). "An efficient transformation procedure enabling long-term storage of competent cells of various yeast genera." Yeast **7**(7): 691-2.

Durfee, T., K. Becherer, et al. (1993). "The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit." Genes Dev **7**(4): 555-69.

Epps, J. L. and S. Tanda (1998). "The Drosophila semushi mutation blocks nuclear import of bicoid during embryogenesis." Curr Biol **8**(23): 1277-80.

Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**(6230): 245-6.

Fremion, F., M. Astier, et al. (1999). "The heterotrimeric protein Go is required for the formation of heart epithelium in Drosophila." J Cell Biol **145**(5): 1063-76.

Garrington, T. P. and G. L. Johnson (1999). "Organization and regulation of mitogen-activated protein kinase signaling pathways." Curr Opin Cell Biol **11**(2): 211-8.

Gietz, R. D., R. H. Schiestl, et al. (1995). "Studies on the transformation of intact yeast cells by the LiAc/SS- DNA/PEG procedure." Yeast **11**(4): 355-60.

Giniger, E., K. Tietje, et al. (1994). "lola encodes a putative transcription factor required for axon growth and guidance in Drosophila." Development **120**(6): 1385-98.

Hanahan, D.; Jessee, J.; Bloom, F.R. (1995). "Techniques for Transformation of E. coli" in: *DNA Cloning*, 2. Edition. D.M Glover and B.D. Hames (Eds.). Oxford University Press. **Vol.1**: 1-36

Harper, J. W., G. R. Adami, et al. (1993). "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin- dependent kinases." Cell **75**(4): 805-16.

Herbst, R., X. Zhang, et al. (1999). "Recruitment of the protein tyrosine phosphatase CSW by DOS is an essential step during signaling by the sevenless receptor tyrosine kinase." Embo J **18**(24): 6950-61.

Hope, I. A. and K. Struhl (1986). "Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast." Cell **46**(6): 885-94.

Hunter, T. (2000). "Signaling--2000 and beyond." Cell **100**(1): 113-27.

Imam, F., D. Sutherland, et al. (1999). "stumps, a Drosophila gene required for fibroblast growth factor (FGF)- directed migrations of tracheal and mesodermal cells." Genetics **152**(1): 307-18.

Inouye, C., N. Dhillon, et al. (1997). "Ste5 RING-H2 domain: role in Ste4-promoted oligomerization for yeast pheromone signaling." Science **278**(5335): 103-6.

James, P., J. Halladay, et al. (1996). "Genomic libraries and a host strain designed for highly efficient two- hybrid selection in yeast." Genetics **144**(4): 1425-36.

Jarecki, J., E. Johnson, et al. (1999). "Oxygen regulation of airway branching in Drosophila is mediated by branchless FGF." Cell **99**(2): 211-20.

Karin, M. and T. Hunter (1995). "Transcriptional control by protein phosphorylation: signal transmission from the cell surface to the nucleus." Curr Biol **5**(7): 747-57.

Keegan, L., G. Gill, et al. (1986). "Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein." Science **231**(4739): 699-704.

Klambt, C., L. Glazer, et al. (1992). "breathless, a Drosophila FGF receptor homolog, is essential for migration of tracheal and specific midline glial cells." Genes Dev **6**(9): 1668-78.

Klint, P., S. Kanda, et al. (1995). "Shc and a novel 89-kDa component couple to the Grb2-Sos complex in fibroblast growth factor-2-stimulated cells." J Biol Chem **270**(40): 23337-44.

Kouhara, H., Y. R. Hadari, et al. (1997). "A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway." Cell **89**(5): 693-702.

Koyano, Y., T. Kawamoto, et al. (1997). "Molecular cloning and characterization of CDEP, a novel human protein containing the ezrin-like domain of the band 4.1 superfamily and the Dbl homology domain of Rho guanine nucleotide exchange factors." Biochem Biophys Res Commun **241**(2): 369-75.

Lee, G. W., F. Melchior, et al. (1998). "Modification of Ran GTPase-activating protein by the small ubiquitin- related modifier SUMO-1 requires Ubc9, an E2-type ubiquitin-conjugating enzyme homologue." J Biol Chem **273**(11): 6503-7.

Leevers, S. J., B. Vanhaesebroeck, et al. (1999). "Signalling through phosphoinositide 3-kinases: the lipids take centre stage." Curr Opin Cell Biol **11**(2): 219-25.

Leevers, S. J., D. Weinkove, et al. (1996). "The Drosophila phosphoinositide 3-kinase Dp110 promotes cell growth." Embo J **15**(23): 6584-94.

Legrain, P., C. Chapon, et al. (1993). "Interactions between PRP9 and SPP91 splicing factors identify a protein complex required in prespliceosome assembly." Genes Dev **7**(7B): 1390-9.

Legrain, P., M. C. Dokhelar, et al. (1994). "Detection of protein-protein interactions using different vectors in the two-hybrid system." Nucleic Acids Res **22**(15): 3241-2.

Lehming N., McGuire S. et al. (1995) "Interactions of a Rel protein with its inhibitor." Proc. Natl. Acad. Sci. USA **92**: 10242-10246

Lindsley D.L., Zimm G.G. (1992). *The Genome of Drosophila melanogaster*. Academic Press, Inc.

Llimargas, M. and J. Casanova (1999). "EGF signalling regulates cell invagination as well as cell migration during formation of tracheal system in Drosophila." Dev Genes Evol **209**(3): 174-9.

Louvet, O., F. Doignon, et al. (1997). "Stable DNA-binding yeast vector allowing high-bait expression for use in the two-hybrid system." Biotechniques **23**(5): 816-8, 820.

Luschnig, S., J. Krauss, et al. (2000). "The Drosophila SHC adaptor protein is required for signaling by a subset of receptor tyrosine kinases." Mol Cell **5**(2): 231-41.

MacDougall, L. K. and M. D. Waterfield (1996). "Receptor signalling: to sevenless, a daughter." Curr Biol **6**(10): 1250-3.

Magee, T. and C. Marshall (1999). "New insights into the interaction of Ras with the plasma membrane." Cell **98**(1): 9-12.

Mahajan, R., C. Delphin, et al. (1997). "A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2." Cell **88**(1): 97-107.

Matunis, M. J., E. Coutavas, et al. (1996). "A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex." J Cell Biol **135**(6 Pt 1): 1457-70.

Matunis, M. J., J. Wu, et al. (1998). "SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex." J Cell Biol **140**(3): 499-509.

Meili, R., C. Ellsworth, et al. (1999). “Chemoattractant-mediated transient activation and membrane localization of Akt/PKB is required for efficient chemotaxis to cAMP in Dictyostelium.” Embo J **18**(8): 2092-105.

Melchior, F. (2000). “SUMO--nonclassical ubiquitin.” Annu Rev Cell Dev Biol **16**: 591-626.

Metzger R.J. and Krasnow M.A. (1999). "Genetic control of Branching Morphogenesis;" Science 1999 **284** (5420): 1635-9

Michelson, A. M., S. Gisselbrecht, et al. (1998). “Heartbroken is a specific downstream mediator of FGF receptor signalling in Drosophila.” Development **125**(22): 4379-89.

Mitchison, T. J. and J. Sedat (1983). “Localization of antigenic determinants in whole Drosophila embryos.” Dev Biol **99**(1): 261-4.

Mohammadi, M., J. Schlessinger, et al. (1996). “Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism.” Cell **86**(4): 577-87.

Ohno, C. K. and M. Petkovich (1993). “FTZ-F1 beta, a novel member of the Drosophila nuclear receptor family.” Mech Dev **40**(1-2): 13-24.

Okada, T., A. Maeda, et al. (2000). “BCAP: the tyrosine kinase substrate that connects B cell receptor to phosphoinositide 3-kinase activation.” Immunity **13**(6): 817-27.

Ong, S. H., G. R. Guy, et al. (2000). “FRS2 proteins recruit intracellular signaling pathways by binding to diverse targets on fibroblast growth factor and nerve growth factor receptors.” Mol Cell Biol **20**(3): 979-89.

O'Shea, J. J., L. D. Notarangelo, et al. (1997). “Advances in the understanding of cytokine signal transduction: the role of Jaks and STATs in immunoregulation and the pathogenesis of immunodeficiency.” J Clin Immunol **17**(6): 431-47.

Parent, C. A., B. J. Blacklock, et al. (1998). “G protein signaling events are activated at the leading edge of chemotactic cells.” Cell **95**(1): 81-91.

Parent, C. A. and P. N. Devreotes (1999). “A cell's sense of direction.” Science **284**(5415): 765-70.

Pawson, T. (1995). “Protein modules and signalling networks.” Nature **373**(6515): 573-80.

Pawson, T. and J. D. Scott (1997). “Signaling through scaffold, anchoring, and adaptor proteins.” Science **278**(5346): 2075-80.

Perkins, L. A., I. Larsen, et al. (1992). “corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso.” Cell **70**(2): 225-36.

Raabe, T., J. Riesgo-Escovar, et al. (1996). “DOS, a novel pleckstrin homology domain-containing protein required for signal transduction between sevenless and Ras1 in Drosophila.” Cell **85**(6): 911-20.

Ramirez-Weber, F. A. and T. B. Kornberg (1999). “Cytonemes: cellular processes that project to the principal signaling center in Drosophila imaginal discs.” Cell **97**(5): 599-607.

Reichman-Fried, M., B. Dickson, et al. (1994). “Elucidation of the role of breathless, a Drosophila FGF receptor homolog, in tracheal cell migration.” Genes Dev **8**(4): 428-39.

Rodrigues, G. A., M. Falasca, et al. (2000). "A novel positive feedback loop mediated by the docking protein Gab1 and phosphatidylinositol 3-kinase in epidermal growth factor receptor signaling." Mol Cell Biol **20**(4): 1448-59.

Rogers, S., R. Wells, et al. (1986). "Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis." Science **234**(4774): 364-8.

Rommel, C. and E. Hafen (1998). "Ras--a versatile cellular switch." Curr Opin Genet Dev **8**(4): 412-8.

Sambrook, J.; Fritsch, E.F.; Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press

Samakovlis, C., N. Hacohen, et al. (1996). "Development of the Drosophila tracheal system occurs by a series of morphologically distinct but genetically coupled branching events." Development **122**(5): 1395-407.

Sanger, F., S. Nicklen, et al. (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-7.

Schlessinger, J. (2000). "Cell signaling by receptor tyrosine kinases." Cell **103**(2): 211-25.

Shishido, E., N. Ono, et al. (1997). "Requirements of DFR1/Heartless, a mesoderm-specific Drosophila FGF- receptor, for the formation of heart, visceral and somatic muscles, and ensheathing of longitudinal axon tracts in CNS." Development **124**(11): 2119-28.

Silver, P. A., L. P. Keegan, et al. (1984). "Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization." Proc Natl Acad Sci U S A **81**(19): 5951-5.

Spradling, A. C., D. Stern, et al. (1999). "The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes." Genetics **153**(1): 135-77.

Staudinger, J., M. Perry, et al. (1993). "Interactions among vertebrate helix-loop-helix proteins in yeast using the two-hybrid system." J Biol Chem **268**(7): 4608-11.

Sternberg, P. W. and J. Alberola-Ila (1998). "Conspiracy theory: RAS and RAF do not act alone." Cell **95**(4): 447-50.

Stokoe, D., L. R. Stephens, et al. (1997). "Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase B." Science **277**(5325): 567-70.

Sutherland, D., C. Samakovlis, et al. (1996). "branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching." Cell **87**(6): 1091-101.

Tautz, D. and C. Pfeifle (1989). "A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback." Chromosoma **98**(2): 81-5.

Vincent, S., R. Wilson, et al. (1998). "The Drosophila protein Dof is specifically required for FGF signaling." Mol Cell **2**(4): 515-25.

Wang, Z. and M. F. Moran (1996). "Requirement for the adapter protein GRB2 in EGF receptor endocytosis." Science **272**(5270): 1935-9.

Wappner, P., L. Gabay, et al. (1997). "Interactions between the EGF receptor and DPP pathways establish distinct cell fates in the tracheal placodes." Development **124**(22): 4707-16.

Wieschau, El; Nüsslein-Volhard, C. (1986). "Looking at embryos" in*: Drosophila - A Practical Approach*. D. B. Roberts (Ed.) IRL Press, Oxford: 199 - 227

Wilson, R. and M. Leptin (2000). "Fibroblast growth factor receptor-dependent morphogenesis of the Drosophila mesoderm." Philos Trans R Soc Lond B Biol Sci **355**(1399): 891-5.

Xu, H., K. W. Lee, et al. (1998). "Novel recognition motif on fibroblast growth factor receptor mediates direct association and activation of SNT adapter proteins." J Biol Chem **273**(29): 17987-90.

Yanisch-Perron, C., J. Vieira, et al. (1985). "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." Gene **33**(1): 103-19.

Appendix I.

Different interaction strengths with different yeast two-hybrid vectors

Identical proteins expressed from different vectors will exhibit different binding behaviour (Legrain et al, 1994). An example from this work is when the insert from the clone i21 (in the pACT vector) was cloned into the pGAD vector and tested for activation of His3 when co-expressed with DB-DofΔC (the same construct used to isolate i21 from the yeast two-hybrid screen). Activation of His3 was strongly reduced, in contrast to the interaction observed in the yeast two-hybrid screen (Figure AI.I)

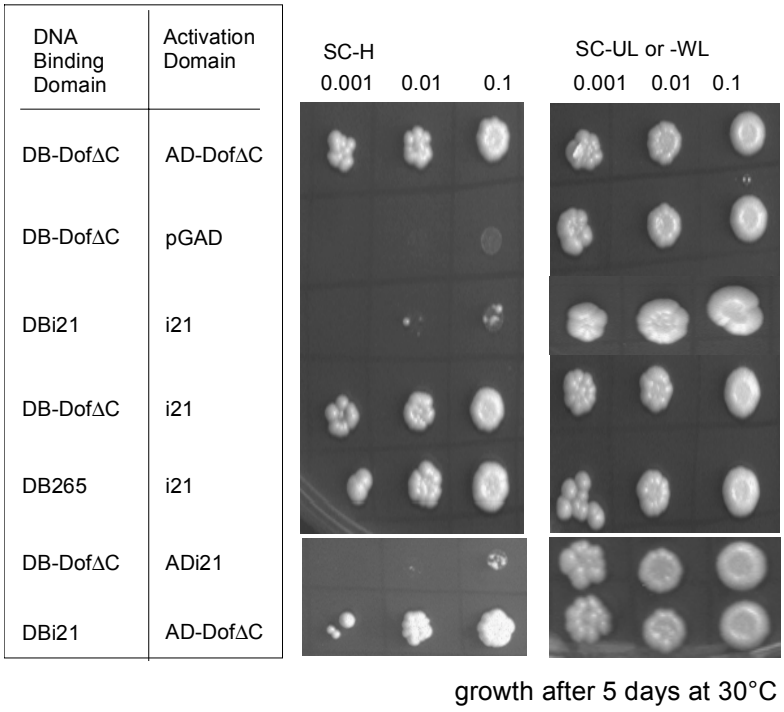


Figure AI.I This figure shows the growth of the construct i21 (in the pACT vector) when co-expressed with DofΔC or full length Dof (construct 256). When the identical sequence is cloned into the pGAD/pGBDU vectors (now called clone ADi21 or DBi21) and co-expressed with DofΔC, the growth rate is diminished. See Figure 2.3 for a description of the dot assay.

Appendix II.

Interaction of candidates with full length Dof and construct 430

All candidates from the yeast two-hybrid screed were tested for activation of the His3 reporter when co-expressed with the full length Dof construct and the deletion construct 430. The results are shown Table A.II.

longest clone	int. DofΔC*	DB 256**	DB 430**
i119	+++	++	*
i59	++	+	+
i14	+++	++++	+
i28	++++	++	++
i238	+	++	+
i56 slight auto	++++	++++	+
i133	++	++++	++++
i150	++++	++++	0
i220	+	+	0
i6	++++	0	+
i19	++++	++++	++++
i159	++++	++++	++++
i199	++	+	++
i7	+++	++++	++++
i163	++++	++	++
i173	+++	0	0
i51	+	0	++++
i239	+	0	+
i31	+	++++	++
i115	+++	0	++
i207	+	0	0
i234 slight auto	+++	++	++
i249	++	++	+
i36	+++	++++	+

longest clone	int. DofΔC*	DB 256**	DB 430**
i94	++	++	0
i70	++	0	++++
i188	+++	0	+
i184	++	0	+
i50	++	++++	++
i164	++		
i17	++	+	0
i155	++	++	+
i226	+++	++	++
i190	++	++++	++
i12	+++	++++	++++
i25 slight auto	+++	++++	++++
i130	++	++	+
i30	+++	++++	++
i118	++	++	++
i8	+++	++++	++++
i136	++++	++++	++++
i169	++	++	+
i175	+	+	+
i176	+	+	+
i182	+++	+	+
i187	+++	0	+
i232	+	0	0

Table AII

* = analysis of growth on SC-ULH selection plates relative to the interaction of AR2.9 (=integrated DofΔC) + ADDofΔC. ++++= faster growth than AR2.9 + AD256, +++ = same growth as AR2.9 + AD256,++= slower growth than int.JH17 + AD256,+ = much slower growth than AR2.9 + AD256, += borderline slow growth**=analysis of growth on dot assay after about 60hrs at 30°C. ++++ = "normal growth" (faster than DBDofΔC + ADDofΔC),++= ½, += ¼, += 1/8, += 1/16, *= 1/32

Appendix III

Analysis of the His3 positive group i119

The group i119 consists of 21 members. Sequence analysis of the 5 prime end revealed that the transcript derives from the predicted gene CG12340. All 21 members code for in-frame fusions with the Gal4 activation domain. The gene encoding the transcript i119 is expressed ubiquitously at early stages but becomes restricted to the nervous system at around stage 12.

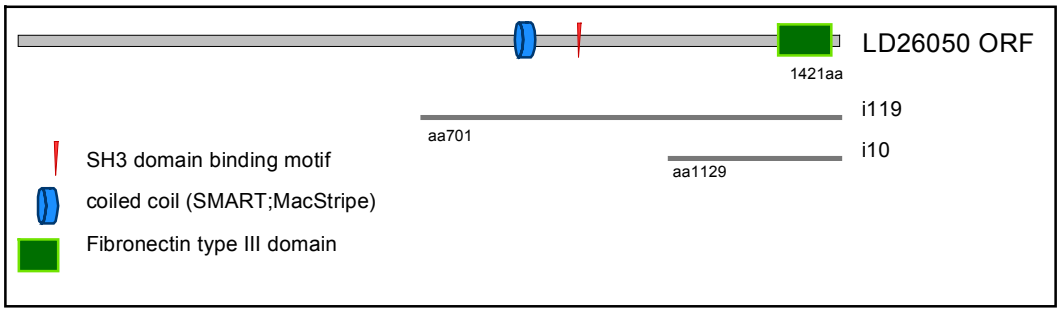


Figure AIII.1 This figure shows the full length protein encoded by the expressed sequence tag cDNA LD26050 aligned with the shortest and longest member of group i119.

The protein sequence of CG12340 does not show significant homology to any known protein. The only obvious motif found in this protein sequence is a Fibronectin type III domain at its C-terminus (PFAM), a coiled coil domain in the middle region of the protein, and a nearby SH3 domain binding motif. The expressed sequence tag LD26050 was identified as a cDNA that is identical to the predicted gene CG12340**.**

Analysis of Embryos that are mutant for the gene encoding i119

CG12340 is cytologically located at 47C1. The deficiencies Df(2R)stan2,b^1pr^1P{ry+t7.2=neoFRT}42D (BL596) and Df(2R)E3363 (BL520) remove the genomic region around 47C1. Tracheal development and mesoderm migration of embryos homozygous for these deficient regions were examined and both deficiencies caused a defect in tracheal development, but heart cell precursors stained with anti-Evenskipped were normal.

Analysis of genes in 47C1

Since the deletion of the genomic region around 47C resulted in defective tracheal development, genes mapping to this region (including the predicted gene CG12340 identified in the yeast two-hybrid screen) were investigated in more detail, since the lack of any one (or more) of them could result in the defect in tracheal development. The genomic region removed by both deficiencies contains not only the gene predicted by CG12340, but several other predicted genes, including some for which mutations already exist. These are *brokenheart* (Fremion et al. 1999), *lola* (Giniger et al. 1994), *pipsqueak* , *stan*, and *diego* (Susan Eaton, personal communication). *lola* (*longitudinals lacking*) is a transcription factor required for the proper development of the central nervous system but not required for the development of the trachea (Giniger et al 1994). *Brokenheart* encodes the alpha subunit of the G_o protein, and is required for the formation of the heart, the visceral musculature and the nervous system in *Drosophila* (Fremion et al 1999). Flies carrying a mutation in Brokenheart were not examined in this work, and it remains to be established whether *Brokenheart*, or its upstream effector G_o has a role in tracheal development. Pipsqueak is a transcription factor that belongs to the posterior group genes. Embryos mutant for *pipsqueak* were not analysed here but it is unlikely that a posterior group gene would exclusively affect tracheal development. The gene *diego* is a novel gene that encodes for a protein that interacts genetically with *stan* (*starry night*), but does not lead to lethality when mutated, indicating that its deletion can not be the sole cause of the tracheal phenotype observed in the deficiency strains BL596 and BL520. Finally, *stan*, a non-homotypic cadherin-like 7-transmembrane protein (involved in cell polarity and axon guidance) is also removed by the

two deficiencies. Two null alleles of stan, *fmi*^{E59} and *fmi*^{E45} (gift from T. Uemura), were analyzed, and both showed a normal tracheal development, indicating that the deletion of *stan* was not causing this phenotype. The last option explored here was to mutate CG12340 to see whether a deletion of this gene is causing the tracheal phenotype observed in the two deficiencies. The mutation of CG12340 was accomplished by P-element mediated excision, the excision was called "ex1.7"and the analysis of this excision is described below. The deletion results in lethality, but tracheal development is normal. Embryos die as pharate adults, some struggling to emerge from their pupal cases. This means that the tracheal phenotype cannot be attributed to the lack of CG12340.

P-element mediated mutagenesis of the predicted gene CG12340

The fly strain EP(2)2024 contains a P-element inserted 76 bases upstream of the first base of genomic sequence corresponding to 5'UTR of LD26050. This insertion does not lead to lethality, meaning that the function of CG12340, the annotated gene coding for LD26050 might not be disrupted by the insertion, or that the disruption of CG12340 is not lethal. In order to determine whether a deletion of CG12340 was the cause of the tracheal phenotype, a deletion of this gene was made using P-element mediated mutagenesis. This was done by mobilizing the P-element in the strain EP(2)2024 and isolating those flies in which the P-element excises imprecisely. All flies in which the P-element had excised could be indentified by their white-eyed phenotype. To determine whether any of these excisions were imprecise and thereby led to lethality, all white-eyed flies were tested for homozygous lethality as well as for non-complemention of the deficiency BL596. Three lethal excisions were isolated and called ex1.7, ex6.1 and ex2.1. They formed two lethal complementation groups. The orientation of these excisions was determined by complementation analysis and PCR. This is described in the next section.

Complementation Analysis of imprecise excisions of EP(2)2024

Imprecise lethal excisions of the neighboring P-element EP(2)2619 (Suzanne Eaton) helped to determine the orientation of the imprecise EP(2)2024 excisions. In addition, the lethal mutations in *starry night* (fmi^{E59} and fmi^{E45}) were used to determine the extent of the 3' excisions. Ex2.1 and ex6.1 excise to the 5' end because they do not complement ex7 and ex 8, but do complement ex10 and ex29. Ex1.7 excises to the 3'end because it complements ex7 and ex8, but does not complement ex10 and ex29.
None of the 3' excisions complement ex212, indicating that the 212 excision removed genomic DNA between EP(2)2619 and EP(2)2024 and that this leads to lethality. The results from the individual complementation crosses are summarized in table AIII.1.

	BL 596	BL 520	1.7	2.1	6.1	ex 10	ex 212	ex 29	ex 7	ex 8	fmi^{E59}	fmi^{E45}	inter se
BL596	L										L	L	L
BL520		L									L	L	L
1.7	L	L	L	V	V	L	V	L	V	V	V	V	L
2.1	L	L	V	L	L	L	V				V	V	L
6.1	L	L	V	L	L	L	V				V	V	L
ex212											(V*)	(V*)	
ex10			L								(L*)	(L*)	L
ex29								L	V	V	(L*)	(L*)	L
ex7								V	L	L	(V*)	(V*)	L
ex8								V	L	L	(V*)	(V*)	L

Table AIII.I Complementation crosses used to determine the direction of the imprecise excisions. ex7 and ex 8 are unidirectional excisions of EP(2)2024 that leave the 3'arm intact.ex10, ex29 and e212 are unidirectional excisions of EP(2)2024 that leave the 5'arm intact (V): S. Eaton, personal communication but unknown which allele of fmi was used for complementation analysis

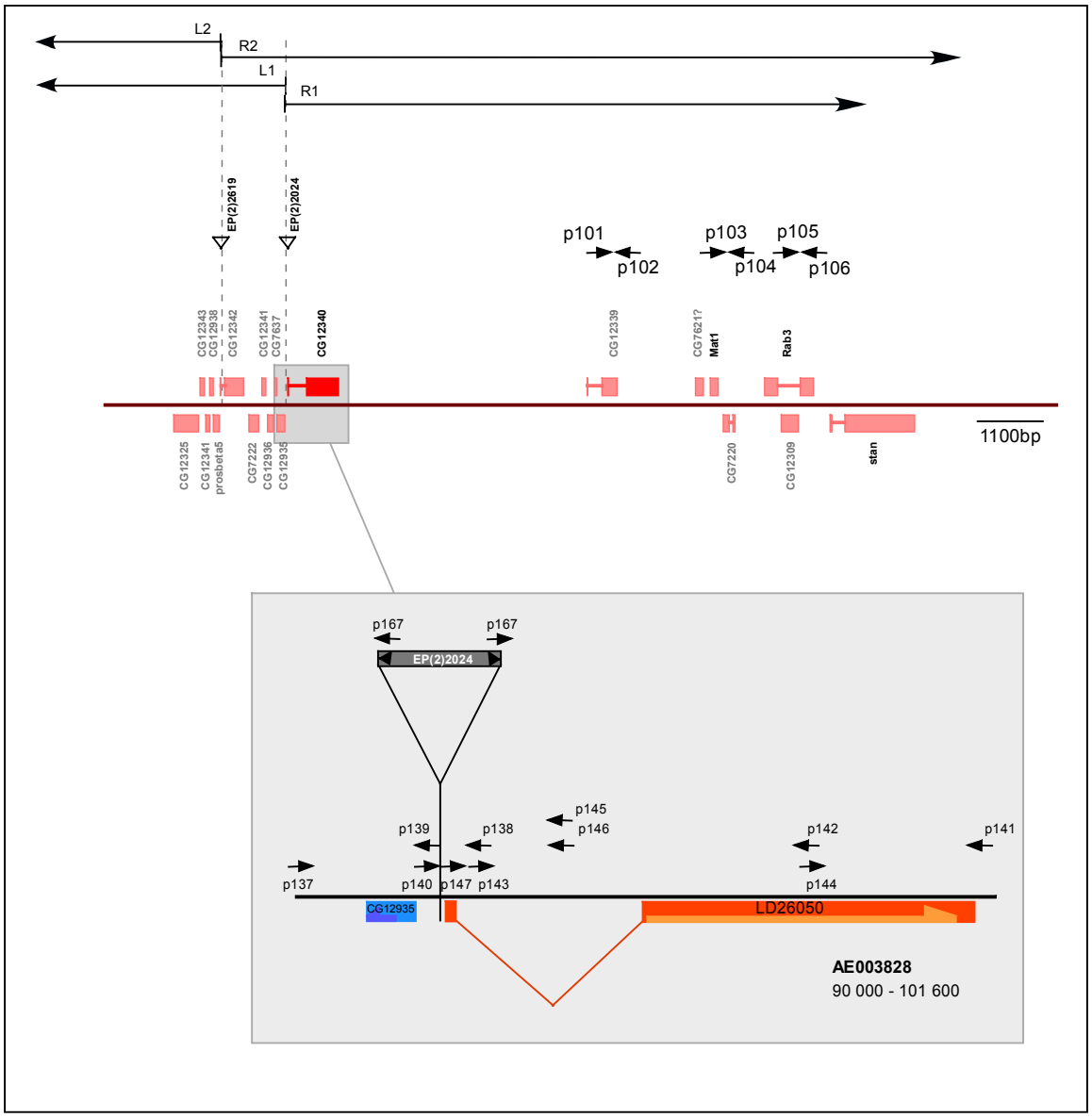


Figure AIII.2 The lethal complementation groups were called L1,L2, R1 and R2 and the direction of excisions is shown in the upper figure. The members of each group are as follows: L2 contains ex7 and ex8. L1 contains ex2.1 and ex6.1. R1 contains ex1.7 and ex10 and R2 contains ex212. Table AIII.2 contains the sequence of primers used for the analysis.

Analysis of excisions using complementation tests and PCR

ex2.1 and ex6.1 contain 5prime excisions

When a P-element excises imprecisely, it is possible that one of the ends is left behind. Therefore, genomic DNA from all three excisions was tested for the presence of one of the P-element ends, which would provide a template for one primer of a PCR product. The gel in Figure AIII.3 shows that the starting strain EP(2)2024 has both inverted repeats that provide a binding site for the P-element primer, giving a product for both sides (lane 3 for both primer pairs). Ex2.1 has an intact 3 prime inverted repeat and genomic region because primers p167 and p138 give the same product as the 3prime arm of EP(2)2024 (lane 5 for primer pair p167 and p138). While the absence of a PCR product does not prove beyond doubt that the genomic region is missing, the data is suggestive that all other excisions lost both ends of the p-element.

pri.#	template	sequence (5'-3')
101	CG12339	aaatccttcggatatacacc
102	CG12339	ccactccgtacaaactgc
103	Mat1	gcaagaccaccaaatacc
104	Mat1	tgaactttctccagctcc
105	Rab3	atacaaaactggcccgtttcc
106	Rab3	ctgtttatgatggtgctgctgc
137	LD26050	ttgctgcagtgcctcctggc
138	LD26050	gagaccatgatgtcgcgca
139	LD26050	ctgatacgcatcgagaagac
140	LD26050	gtcttctcgatgcgtatcag
141	LD26050	actactcaaggatgatgaac
142	LD26050	taaccatcatcgcctcgacg
143	LD26050	tgcgcgacatcatggtctc
144	LD26050	cgtcgaggcgatgatggtta
145	LD26050	ggaaacaggcagctagagcg
146	LD26050	ttgttggtcgcgagcggcag
147	LD26050	ctgctcttcggaaccagttg

Table AIII.2 This table lists the primers used for the analysis of the genomic region around CG12340.

ex1.7 contains a 3 prime excision

Complementation analysis suggested that ex1.7 contained an intact 5 prime genomic region (if the imprecise excision has deleted part of the 5prime region, it in any case does not lead to lethality) and that the 3 prime deletion did not extend as far as the gene *starry night*. PCR with inverted repeats suggested that neither inverted repeat was left behind by the excision (Figure AIII.3). In order to determine the extent of the 3 prime excision, primer pairs from the three closest genes on the 3prime side of EP(2)2024 were used to identify the presence or absence of a genomic sequence.

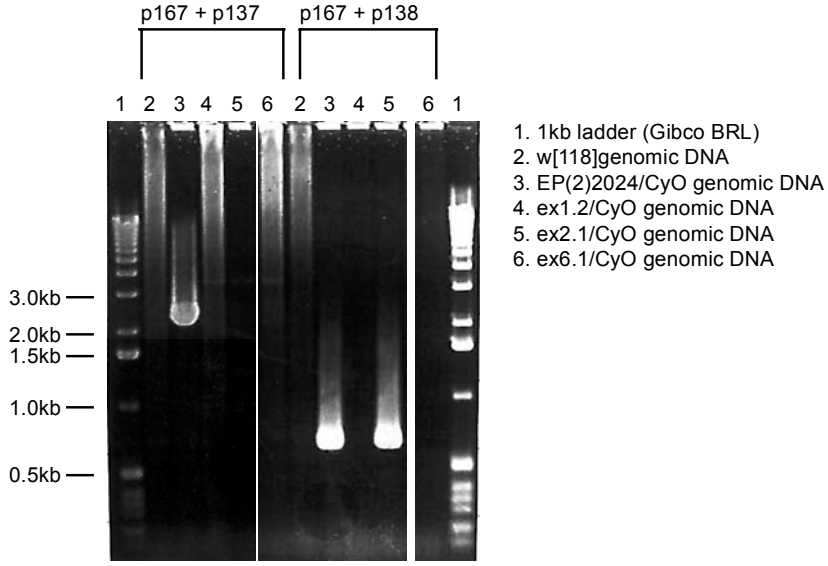


Figure AIII.3. This figure shows the PCR products from PCR reactions on genomic DNA of the excision strains ex1.2/CyO, ex.2.1/ CyO and ex. 6.1/ CyO. Note that ex1.2/ CyO is the same strains as ex1.7/ CyO. The positions of the primers are shown in Figure AIII.2.

For this experiment, homozygous ex1.7/ex1.7 or heterozygous ex1.7/CyO^{actGFP} larvae were selected (by the absence or presence of GFP) and pooled into groups of four. Genomic DNA was prepared from each pool of four larvae. Fig AIII.4 shows that in wild type embryos, all 4 genes on the 3 prime side are present. However, in ex1.7/ex1.7 homozygotes, CG12340 is missing whereas the other 3 are still present. Thus the excision removes the gene encoding LD26050 but does not extend as far as the predicted gene CG12339.

Phenotype of ex1.7

mAB2a12 staining of ex1.7/CyO x ex1.7/CyO collections showed that the trachea are normal and indicated that the tracheal phenotype observed in Df520 and Df596 is not due to the lack of CG12340. In fact, ex1.7/ex1/7 homozygotes only die as pharate adults. Out of 32 homozygotes (identified by GFP^- larvae from w;ex1.7/CyO^{actGFP} parents), two even emerged out of their pupal case, but couldn't stand up. Others struggled to emerge from their pupal cases, but failed to emerge.

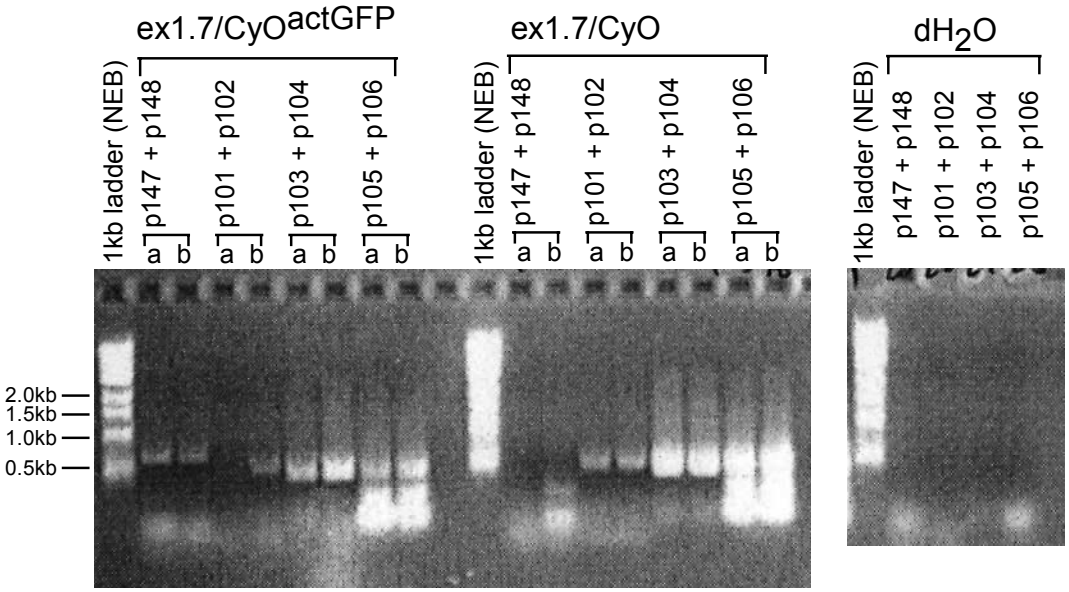


Figure AIII.4 This figure shows the products of PCR reactions on genomic DNA isolated from larvae. The position of the primers is shown in Figure AIII.3

Full length CG12340 interaction with DofΔC

Since only a third of the CG12340 transcript was isolated in the yeast two-hybrid screen, the full length transcript was cloned into the yeast two-hybrid vectors. The clone 271 contains a fusion of the full length protein fused to the activation domain, the clone 272 contains a fusion of the full length protein fused to the DNA binding domain. The full length clone only interacts weakly with DofΔC in both directions. It should be pointed out that the interaction was observed with i119 in the pACT2 vector, whereas the full length fusion constructs were cloned into pGAD and pGBDU (see Appendix I.).
The clones used to generate the construct 271 (DNA-binding fusion in the vector pGBD) and construct 272 (activation domain fusion in the vector pGAD) are shown in Table AIII.2.

construct #	cloning method
212	LD26050 (in pOT2 RI/XhoI)
226	pGBD-C3 NotI- BamHI digest, Klenow fill, NcoI linker NEB 1151 addition
227	pGAD-C3 NotI-BamHI digest, Klenow fill, NcoI linker NEB 1151 addition
231	A226 NotI, klenow fill, add AvrII linker
232	A227 NotI, klenow fill, add AvrII linker
265	SmaI-SpeI digested PCR fragment *** fromORF LD26050 into SmaI-AvrII A231
266	SmaI-SpeI digested PCR fragment*** from ORF LD26050 into SmaI-AvrII A232
268	A265 shortened with NarI
269	A266 shortened with NarI
271	NcoI-SacII A268 + NcoI-SacII A212
272	NcoI-SacII A269 + NcoI-SacII A212

Table AIII.2 This table describes the intermediate clones used to generate constructs 271 and 272. ***primers for PCR: p124 (catcccgggcggccgcatgatgggagtaaaccagac; SmaI site and NotI site in front of ATG) and p125 (gactagtacgaccaggtcttgggaaca; has SpeI site)

Appendix IV: Constructs

Table IV.1 describes the clones used in this project .

NAME OF CONSTRUCT	CLONING METHOD
DB-DofΔC	2.4kb NotI/BglII partial of 1f11 (Dof in pNB40) into BamHI/NotI pGBDU-C2 NotI (from J.Hancke)
int. JH17	DB-DofΔC XhoI/NotI cloned into A146
AD-DofΔC	insert from DB-DofΔC into pGAD-C_2NotI via EcoR1/Not1 (from J.Hancke)
DB409C	DB-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of Dof D233-449*
AD409C	AD-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of Dof D233-449*
DB403C	DB-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of DofD361-449*
AD403C	AD-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of DofD361-449*
DB430C	DB-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of DofD233-364*
AD430C	AD-DofΔC BglII-ClaI gap repair** with NcoI-XhoI of DofD233-364*
AD341	AD-DofΔC Ehe-ClaI gap repair** with Sfi I/XhoI of DofDE531*
AD256	JH14 insert into AD-DofΔC via NotI-EheI
DB256	JH14 insert into DB-DofΔC via NotI-EheI
DB426	DofDS89/E802* NcoI-XbaI into A231 NcoI-AvrII
AD426	DofDS89/E802* NcoI-XbaI into A232 NcoI-AvrII
DB427	DofDS168/E802* NcoI-XbaI into A231 NcoI-AvrII
AD427	DofDS168/E802* NcoI-XbaI into A232 NcoI-AvrII
DB i21	i21 NcoI-SpeI into A231NcoI-AvrII
AD i21	i21 NcoI-SpeI into A232NcoI-AvrII
DB352	DofDS89* NcoI-SpeI into A231NcoI-AvrII
AD352	DofDS89* NcoI-SpeI into A232NcoI-AvrII
DB353	DofDS168* NcoI-SpeI into A231NcoI-AvrII
AD353	DofDS168* NcoI-SpeI into A232NcoI-AvrII
DB354	DofDS227* NcoI-SpeI into A231NcoI-AvrII
AD354	DofDS227* NcoI-SpeI into A232NcoI-AvrII
DBi21-cc	i21 XhoI, fill, add NcoI linker, clone NcoI-SpeI fragment into A#231 NcoI-AvrII
ADi21-cc	i21 XhoI, fill, add NcoI linker, clone NcoI-SpeI fragment into A#232 NcoI-AvrII
AD430	AD256 BglII-ClaI gap repair** with DofD233-364* NcoI-XhoI
DB430	JH14 EheI-NotI insert into DB430C EheI-NotI
DB409	JH14 EheI-NotI insert into DB409C EheI-NotI
AD409	JH14 EheI-NotI insert into DB409C EheI-NotI
AD403	JH14 EheI-NotI into AD403C EheI-NotI
DB403	JH14 EheI-NotI insert into DB403C EheI-NotI
AD405	DofDE453* BglII-XbaI into A261 BglII-AvrII
DB405	DofDE453* BglII-XbaI into A262 BglII-AvrII
AD404	DofDE531* BglII-XbaI into A261 BglII-AvrII
DB404	DofDE531* BglII-XbaI into A262 BglII-AvrII
JH15	JH11 EcoRI-NotI into pGBDc1NotI EcoRI-Not (J.Hancke)
JH11	Clone BglII-Stu>NotI fragment from 1f11 (Dof) into pGAD C2 Not1 BamHI-NotI (from J.Hancke)
JH14	JH10 EcoRI-Not into pGBDc1NotI EcoRI-NotI (from J.Hancke)
JH10	2.8kb BglII-NotI from 1f11 (Dof) into pGADc2NotI BamHI-NotI (from J.Hancke)
JH16	JH13 EcoRI-NotI into pGBDc1-NotI EcoRI-Not (from J.Hancke)
JH13	JH10 XhoI, blunt, NotI linker,close. BglII-XhoI is 1.58kb (from J.Hancke)
pGAD-C(x)	James et al. 1996
pGBD-C1	James et al. 1996
pGBDU-C1	James et al. 1996
pODB8	Louvet et al. 1997
pGAD-C1 NotI	pGAD-C1 BglII digest, fill-in, NotI linker addition (J.Hancke)
pGAD-C2 NotI	pGAD-C2 BglII digest, fill-in, NotI linker addition (J.Hancke)
pGAD-C3 NotI	pGAD-C3 BglII digest, fill-in, NotI linker addition (J.Hancke)
pGBD-C3 NotI	pGBD-C3 BglII digest, fill-in, NotI linker addition (J.Hancke)
pGBDU-C2 NotI	pGBDU-C2 BglII digest, fill-in, NotI linker addition (J.Hancke)
pUASTλBtl	gift from D. Montel. Note: the intron in this Breathless cDNA was not spliced out.
pAlter-1	Promega
A121	pODB8 BamHI>NotI
A146	int. pGBDU-C2 NotI (2μ removed by discarding AatII-SnaBI fragment from pGBDU-NotI, blunting, then religating)
A226	pGBD-C3 NotI digested with BamHI, filled, NcoI linker addition (NEB 1151)
A227	pGAD-C3 NotI digested with BamHI, filled, NcoI linker addition (NEB 1151)

NAME OF CONSTRUCT	CLONING METHOD
A231	A226 NotI, klenow fill, add AvrII linker (NEB#1123)
A232	A227 NotI, klenow fill, add AvrII linker (NEB#1123)
A261	AD256 NotI, fill in, add AvrII linker (Neb1123)
A262	DB256 NotI, fill in, add AvrII linker (Neb1123)
DB-Btl-K	cyt btl in pACT2ΔframeI: EcoRI-SalI AD-Btl-K into EcoRI-XhoI A296
AD-Btl-K	cyt btl in pODB8 (EcoRI-BglII A275 into EcoRI-BamHI pODB8)
A296	NcoI digest A295, fill, close)
A295	BamHI digest pACT2, fill, close
A275	XbaI A#273 insert into XbaI A#208.1
A273	Breathless cDNA without intron. Made by looping the intron out of the Breathless sequence in A267 using the Altered Sites II in vitro Mutagenesis System (Promega). The mutagenesis primer used was 5'gcttcaagacggtgcattcatcgc 3'
A267	pUASTλBtl NotI into NotI A258
A258	pAlter-1 SmaI, add NotI linker (NEB#1126) close
A208.1	BglII-EcoRI digested PCR product from pUASTλBtl (primers 5'agagaattcgtggagtatgctccacacg3' and 5'agagagatctaggtgtactgatatctaagc3') into BglII-EcoRI pGAD-C2 . Note: This cDNA contains intron!
DB-Htl	BglII-XbaI of Htl cDNA (B. Shilo, BglII linker NEB#1052 introduced at XbaI site) into BamHI-XbaI of A#209
A307	i8 BamHI digest. Fill (frameshift#1)
A308	A307 NcoI digest. Fill (frameshift#2) Sequenced.
DB-Δs.site#3	BsshII-EheI A288 into BsshII-EheI A304
AD-Δs.site#3	BsshII-EheI A288 into BsshII-EheI A306
A306	AD-DofΔC BglII-SfiI. Discard 400bp fragment. blunt. close
A304	DB-DofΔC BglII-SfiI. Discard 400bp fragment. blunt. close
A288	1f11in pAlter containing mutated sumoylation site#3 in. Made by introducting an Arginine to Lysine (aag becomes aga) at third SUMOlation site of Dof using the Altered Sites II in vitro Mutagenesis System (Promega). The mutagenesis primer used was 5'gacccatcagatgcgaaagtc 3'
1f11in pAlter	1f11+5'FLAGTAG in pNB40 into pAlter via EcoRI-SphI . Construct from R. Wilson

Table AIV.1 This table describes the constructs used in this work. Constructs discussed in the results section are in bold type. All others are intermediate constructs used for cloning. **gap repair was performed as described by N. Lehming et al. (1995). *All cDNAs with deletions in Dof were obtained from R. Wilson

Abbreviations

Ade	Adenine
ATP	Adenosine Triphosphate
bp	base pairs
BSA	Bovine Serum Albumin
cDNA	complementary DNA
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetracetic Acid
g	gram
GST	Glutathione-S-transferase
His	Histidine
k	kilo
L	Leucine
LacZ	beta-galactosidase
M	mol per liter
m	milli
µ	micro
mRNA	messenger RNA
n	nano
NGS	Normal Goat Serum
PCR	polymerase chain reaction
RNA	ribonucleic acid
SC-U	Synthetic Complete medium lacking Uracil
SC-UL	Synthetic Complete medium lacking Uracil and Leucine
SC-ULH	Synthetic Complete medium lacking Uracil, Leucine and Histidine
SC-H	Synthetic Complete medium lacking Histidine
SC-A	Synthetic Complete medium lacking Adenine
X-Gal	5-Bromo-4chloro-3-indoyl-ß-D-galactoside

Abstract

Downstream of FGFR (Dof) is an adaptor protein that is required specifically for Fibroblast Growth Factor (FGF) signal transduction in *Drosophila melanogaster*. Although Dof has been shown to function upstream of the Ras MAPK pathway it is unknown how Dof transduces the FGF signal from the FGF receptor to the Ras MAPK pathway. It is also unclear whether Dof is required for the transduction of additional downstream signaling pathways that might be activated by the FGF receptors. One way to resolve this question is to identify the molecules required for Dof's function. In this project, a yeast two-hybrid screen was conducted to identify interacting partners of Dof. Eight groups and three single clones tested positive in three different assays employing different reporters with different promoters, some more stringent than others. The most stringent assay eliminated about 70% of the candidates. Those that remained interactors in a less stringent assay might represent "weak" interactions and were therefore also investigated. Three candidates that tested positive for all selection assays provide clues to how Dof may be functioning *in vivo*. First, the Drosophila FGF receptor Heartless was isolated. Subsequent work showed that the second Drosophila FGF receptor, Breathless, also interacts with Dof in the yeast two-hybrid system. These results suggest that Dof might function by binding directly to the receptor. The second candidate of interest is Ubc9, which conjugates SUMO (small ubiquitin-related modifier) moieties onto target proteins. SUMO has been found to regulate protein-protein interactions and/or subcellular localization of target proteins. Dof has four putative SUMOlation sites and when the third putative SUMOlation site was mutated, this completely abolished binding of Dof to both FGFRs in the yeast two-hybrid system, indicating that the binding affinity of Dof to the FGFRs might be regulated by post-translational SUMOlation. The third candidate of interest from the screen was Dof itself, revealing that Dof can self-associate. In order to characterize this homodimerization in more detail, the domains in Dof that are responsible for self-association were mapped by co-expressing a panel of Dof deletion constructs against each other in the yeast-two hybrid system. This experiment revealed that a region of the N-terminus can bind both to itself and to a region in the C-terminus of the molecule.

The yeast two-hybrid partners reveal that Dof can potentially receive inputs at various levels: It could be receiving a signal directly from the FGF receptors and this might be dependent on SUMOlation. In addition, Dof could be using dimerization to channel other signaling molecules into a specific pathway and thereby transduce the FGF signal.

Zusammenfassung

Der Faktor Downstream of FGFR (Dof) ist ein Adaptorprotein, das spezifisch für den FGF Signaltransduktionweg in *Drosophila melanogaster* benötigt wird. Genetische Experimente deuten darauf hin, daß Dof stromaufwärts des Ras MAPK Signaltransduktionweges arbeitet. Sein genauer Platz im Signalweg und der eigentliche Mechanismus, wie Dof das Signal vom aktivierten Rezeptor zur MAPK Kaskade weiterleitet, ist allerdings noch nicht bekannt. Ebenfalls unklar ist, ob Dof auch für die Signaltransduktion in anderen, vom FGF-Rezeptor aktivierten Signalwegen benötigt wird. Eine Möglichkeit um diese Fragen zu klären, ist die Identifizierung und Charakterisierung von Faktoren, die für die Funktionen von Dof gebraucht werden. Um solche Proteine zu finden wurde in der vorliegenden Arbeit ein yeast two-hybrid Experiment durchgeführt. Dieses Experiment wurde mit verschiedenen Reporterkonstrukten mit unterschiedlich stringenten Promotoren durchgeführt. Der stringenteste Promotor reduzierte die Kandidatenanzahl um 70%. Die Kandidaten, die nur unter weniger stringenten Bedingungen Effekte zeigten, wurden als schwache Interaktoren klassifiziert und ebenfalls untersucht. Insgesamt acht Gruppen und drei einzelne Klone zeigten Interaktion in den drei obengenannten Testsystemen. Schließlich wurden drei Kandidaten ermittelt, die einen Hinweis auf die Wirkungsweise von Dof im *Drosophila* Embryo geben könnten.

Ein erster Kandidat ist einer der Rezeptoren des FGF Signalweges, Heartless. Weiterführende Arbeiten konnten zeigen, dass auch der zweite Rezeptor, Breathless, im yeast two-hybrid Experiment mit Dof interagiert. Diese Ergebnisse deuten darauf hin, daß Dof direkt an den Rezeptor bindet und so die Signalweiterleitung steuert.

Ein zweiter Faktor, der im Rahmen dieser Arbeit identifiziert wurde, ist das Protein Ubc9 (Ubiquitin Conjugating Enzyme). Ubc9 kann einen Ubiquitin verwandten Faktor namens SUMO (small ubiquitin-related modifier) auf Zielproteine übertragen. Es ist bekannt, daß durch solche Modifikationen Proteininteraktionen und/oder Proteinlokalisierungen innerhalb der Zelle, reguliert werden. Das Dof Protein besitzt vier Stellen an denen es möglicherweise mit SUMO verbunden werden könnte. Die Analyse der einzelnen Erkennungssequenzen wurde wieder im yeast two-hybrid Ansatz durchgeführt. Diese ergab, dass das Dof mit einer Mutation in der dritten SUMO Verknüpfungsstelle nicht mehr mit den beiden FGF Rezeptoren interagieren kann. Die Bindung von Dof an die Rezeptoren könnte also durch die Ankopplung von SUMO reguliert bzw. ermöglicht werden.

Der dritte Kandidat ist Dof selber, was vermuten lässt, dass Dof Homodimere bilden kann. Um diese Dimerisierung näher zu charakterisiern wurden die dafür notwendigen Proteindomänen bestimmt. Hierfür wurden, wieder im yeast two-hybrid System, mehrer Deletionskonstrukte in verschiedenen Kombinationen exprimiert. Diese Analyse zeigt, dass eine bestimmte Region im N-Terminus des Proteins mit sich selber und mit einer c-terminalen Region interagieren kann.

Die Identifizierung der oben genannten Kandidaten im Rahmen dieser Arbeit, zeigt wie vielfältig die Wirkungsweise und Regulation von Dof möglicherweise ist. Eine Regulation ist, wie oben gezeigt, sowohl auf der Ebene des Rezeptors, als auch durch Modifikationen von Dof selber, entweder durch Dimerisierung oder Verknüpfung mit SUMO,denkbar.

Ich versichere, daß ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen - die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie noch nicht veröffentlicht worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der geltenden Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maria Leptin betreut worden.

Teilveröffentlichungen liegen nicht vor.

Alysia Battersby

Acknowledgements

The work presented in this thesis was done at the Institut für Genetik der Universität zu Köln under the supervision of Prof. Dr. Maria Leptin.

I would like to thank Maria for creating a challenging and inspiring scientific environment and for providing endless support, especially at the end of this project. My thanks also goes to members of the Leptin , Sprenger, Klein and Roth labs for many helpful discussions. In particular, I am grateful to Maithreyi Narasimha for her comments on this thesis, Ruth Grosskortenhaus for help with translations, and members of the Dof minigroup, in particular Robert Wilson and Min Yan Zhu for their helpful suggestions throughout the project. I would also like to thank Lisa Vogelsang for keeping the lab running smoothely and Heidi Thelen for running endless sequencing gels. Last, but not least, I would like to thank Juliane Hancke for her technical assistance in the final stages of the project.

My biggest thanks goes to my parents and Mike for giving me strength when I needed it most.

Lebenslauf

Alysia Battersby geb. Runde	Staatsangehörigkeit: deutsch verheiratet mit Michael Battersby
Erster Wohnsitz:	Wichterichstr. 19 50937 Köln
06.10.1970	geboren in Pinelands – Cape, Südafrika Eltern: Alfons Runde und Pamela Runde geb. Lord
1976 – 1986	Grundschule und Hauptschule
1986 – 1989	Gymnasium in Mettingen, Abitur im Mai 1989
1990 – 1994	Bachelor of Arts, University of Pennsylvania, Pennsylvania, U.S.A.
1994 – 1996	Biologie-Studium an der Universität zu Köln
1996 – 1997	Diplomarbeit bei Frau Prof. Dr. Maria Leptin am Institut für Genetik der Universität zu Köln Titel:“Analysis of a novel gene expressed in the early mesoderm of the Drosophila embryo”
September 1997 – Mai 2001	Doktorarbeit bei Frau Prof. Dr. Maria Leptin am Institut für Genetik der Universität zu Köln Titel: “Identification of molecules that interact with the adaptor protein Dof”