FOR THE RECORD

Identification of a Frizzled-like cysteine rich domain in the extracellular region of developmental receptor tyrosine kinases

JOSE SALDANHA,¹ JUSWINDER SINGH,² AND DARUKA MAHADEVAN³

¹Division of Mathematical Biology, National Institute for Medical Research, London NW7 1AA, United Kingdom

²Biogen, Inc., 14 Cambridge Center, Cambridge, Massachusetts 02142

³Birkbeck College, University of London, London WC1E 7HX, United Kingdom

(RECEIVED May 22, 1998; ACCEPTED June 1, 1998)

Abstract: In *Drosophila*, members of the Frizzled family of tissue-polarity genes encode proteins that appear to function as cell-surface receptors for Wnts. The Frizzled genes belong to the seven transmembrane class of receptors (7TMR) and have on their extracellular region a cysteine-rich domain that has been implicated as the Wnt binding domain. This region has a characteristic spacing of ten cysteines, which has also been identified in FrzB (a secreted antagonist of Wnt signaling) and Smoothed (another 7TMR, which is involved in suppression of the hedgehog pathway). We have identified, using BLAST, sequence similarity between the cysteine-rich domain of Frizzled and several receptor tyrosine kinases, which have roles in development. These include the muscle-specific receptor tyrosine kinase (MuSK), the neuronal specific kinase (NSK2), and ROR1 and ROR2. At present, the ligands for these developmental tyrosine kinases are unknown. Our results suggest that Wnt-like ligands may bind to these developmental tyrosine kinases

Keywords: cysteine-rich domains; developmental tyrosine kinases; Frizzled; Wingless

The interaction of Wingless (Wnt) with its receptor Frizzled (fz) activates specific intracellular signaling pathways, implicated in embryogenesis and oncogenesis. Wnts are a family of cysteine rich, highly basic glycoprotein ligands implicated in cell-lineage decisions in *Caenorhabditis elegans*, in embryonic and adult pattern formation in *Drosophila*, in axis formation and dorsal-ventral polarity determination in *Xenopus* embryos, and in central nervous system development and oncogenesis in mice and humans (Bhanot et al., 1996). The Frizzled genes named after the *Drosophila* tissue polarity gene frizzled, belong to the seven transmembrane receptor (7TMR) family but additionally contain an extracellular cysteine-

rich Wnt binding domain (CRD). The sequence following the CRD is of variable length among the frizzled receptors and in *Drosophila* frizzled 2 (Dfz2R) is glycine and serine rich (Bhanot et al., 1996). This domain may serve as an extended flexible linker between the CRD and the membrane spanning domain. Frizzled joins a class of 7TMRs, including the Luteinizing/Chorionic Gonadotrophin hormone (LH/CG) receptor, follicular stimulating hormone (FSH) receptor, thyroid stimulating hormone (TSH) receptor, and the calcium ion receptor, which all contain large extracellular ligand binding domains (Dale, 1998).

The CRD found in the Frizzled family has been identified in the genes of the frizbee (Frzb) family (Leyns et al., 1997; Wang et al., 1997), a splice variant of collagen XVIII (long form) (Rehn & Pihlajaniemi, 1995), the Smoothed receptor (SmoR) (Alcedo et al., 1996), the N-terminal region of Carboxypeptidase Z (CPZ) (Song & Fricker, 1997) and the secreted apoptosis-related proteins 1, 2, and 3 (Melkonyan et al., 1997), which appear to be Frzb-like molecules. The CRD of the frizzled receptor is about 130 amino acids in length and is located at the N-terminus of the receptor. It contains 10 cysteines, which are largely conserved among the frizzled receptors and related sequences. The conservation of the 10 cysteines throughout evolution suggests that this domain contains a number of structurally important disulfide bridges.

Sequence analysis: A position specific iterated Blast (PSI-Blast) (Altschul et al., 1997) search was run on the nonredundant database at the National Center for Biotechnology Information (NIH, Bethesda, Maryland) using the CRD from Dfz2R as a probe. PSI-Blast produces position-specific score matrices from the gapped Blast2.0 output and employs them to iterate the database search, thus picking out more distant homologues. The search was performed without a sequence filter and the expect value for inclusion of sequences in each iteration was kept at 0.01. Other parameters defaulted. The search converged at the fourth iteration, and the results contained not only the various related sequences detailed above but also the muscle specific tyrosine kinase receptor MuSK from rat (Valenzuela et al., 1995) with an expect value (the prob-

Reprint requests to: Juswinder Singh, Biogen, Inc., Department of Drug Design and Evaluation, 12 Cambridge Center, Cambridge, Massachusetts 02142; e-mail: juswinder_singh@biogen.com.

Fig. 1. Sequence alignment of CRDs from members of the frizzled and receptor tyrosine kinase families. The sequence names and PID numbers are: dfz2r, *Drosophila* frizzled 2 receptor (g1518051); fz_human, human frizzled homologue (g1151252); fz_rat, rat frizzled homologue (g310115); fz_mouse, mouse frizzled homologue (g1151260); fz_gal, chicken frizzled-7 (g2655276); frzb_human, human frizbee (g1620541); frzb_bos, bovine frizbee (g1620539); frzb_xen, *Xenopus laevis* frizbee (g1854030); apop_human, human secreted apoptosis related protein 3 (g2415419); cpz_human, human carboxypeptidase Z (g2160714); smoo_human, human smoothened receptor (g1813876); coll_mouse, mouse collagen alpha 1(XVIII) chain (g1363109); musk_rat, rat muscle-specific tyrosine kinase receptor (g1015392); mlk_mouse, mouse muscle localized kinase 1 (g1017827); musk_human, human muscle-specific tyrosine kinase receptor (g2253312); nsk2_mouse, mouse Nsk2 protein (g2137344); rtk_per, Pacific electric ray probably protein-tyrosine kinase RTK (g422541); ror1_human, human probable protein tyrosine kinase receptor ror1 (g346351); nrtk_dros, *Drosophila* neurospecific receptor tyrosine kinase (g2116554). The largely conserved cysteines have been highlighted.

ability of finding the alignment by chance) of 3e-20, the muscle localized kinase 2 from mouse (3e-20) (Caruso et al., 1995), the human muscle specific tyrosine kinase receptor (3e-20) (Valenzuela et al., 1995), the muscle localized kinase 1 and 2 from mouse

(1e-19) (Valenzuela et al., 1995), mouse NSK2 (4e-19) (Ganju et al., 1995), and the probable receptor tyrosine kinase (RTK) from *Torpedo californica* (2e-11) (Jennings et al., 1993). Sequences below the expect value threshold of 0.01, but with plausible align-

ments, were the probable human RTKs ROR1 and ROR2 (0.029 and 0.25, respectively) (Masiakowski & Carroll, 1992) and the neurospecific RTK from *Drosophila* (0.97) (Oishi et al., 1997). Sequence alignment of representative sequence segments from the frizzled and RTK families was performed with ClustalX (Thompson et al., 1994) plus some minor manual adjustments (Fig. 1). It is clear from the figure that the pattern of 10 cysteine residues is largely conserved, the exceptions being in the second cysteine, which is methionine in NSK2, and the sixth cysteine, which is serine in Smoothed. The basic residue (R/K) immediately following the sixth cysteine is also conserved, except in the neurospecific RTK where there is an F in that position. There are also conserved patterns of hydrophobic residues. Thus, we illustrate that part of the extracellular domain of receptor tyrosine kinases (RTKs) has the conserved pattern of 10 cysteines observed for the frizzled family of receptors.

Identification of *crd* in receptor-tyrosine kinases: RTK-mediated signaling is utilized extensively to mediate cell-cell interactions during normal development of multicellular organisms (Ganju et al., 1995; Valenzuela et al., 1995). The developmental RTKs mentioned above are expressed during embryogenesis in early myotomes, and during muscle and neural development. Subsequently, their expression is markedly diminished except for MuSK expression, which is increased at the neuromuscular junction and during stressful stimuli (e.g., injury). Most growth factor receptors are rapidly down-regulated following cessation of the proliferative phase of myoblast development, but MuSK expression remains high throughout early muscle development in vivo and increases as myoblasts differentiate during culture in vitro (Valenzuela et al., 1995). Thus, in contrast to growth factors, the putative ligand for MuSK may continue to serve a role during later stages of differentiation. The distribution of NSK2 transcripts during mouse development implicate functions for this RTK in skeletal myogenesis, neural development, and mesenchymal-epithelial interactions during fetal organogenesis (Ganju et al., 1995). Since the frizzled receptors are also expressed during embryogenesis, it is conceivable that Wnts or Wnt-like ligands interact with CRDs of the RTKs. The extracellular domains of RTKs contain a number of different domain types such as immunoglobulin-like domains (Fibroblast growth factor receptor, platelet derived growth factor receptor), cysteine containing domains (epidermal growth factor receptor, insulin receptor), and fibronectin-like domains (Lewis et al., 1996). The presence of frizzled-like CRDs attached to developmental RTKs suggests a mechanism of how RTKs may have been utilized to provide developmental signals.

The extracellular domain of ROR1 and ROR2, and neurospecific RTK (close relatives of the Trk RTK family) are known to contain a cysteine-rich domain (127 amino acids containing 10 cysteines) followed by a kringle domain (80 amino acids) (Masiakowski & Carroll, 1992; Oishi et al., 1997). This cysteine-rich domain coincides with the CRD of the frizzled receptors according to our alignment (Fig. 2). The Torpedo RTK was postulated to contain four Ig-like domains (each of 50 to 55 amino acids), with a cysteine-rich domain (78 amino acids), and a kringle domain (80 amino acids) either side of the fourth Ig-like domain (Jennings et al., 1993). NSK2 and MuSK are similar to the Torpedo RTK except that they do not have a kringle domain (Ganju et al., 1995; Valenzuela et al., 1995). The CRD of the frizzled receptors encompasses a region in these sequences including the cysteine-rich

Fig. 2. Schematic representation of the CRDs within the frizzled family and developmental receptor tyrosine kinases.

domain and the fourth Ig-like domain suggesting that the latter assignment may be incorrect.

The RORs appear to be distant relatives of MuSK and Torpedo RTK and homology in their extracellular domain suggests that they may recognize structurally related ligands. It is postulated that developmental RTKs may respond to multiple ligands during development. The kringle domain (found in hepatocyte growth factor and thrombin) may interact via protein-protein interactions, the frizzled-like CRD may bind Wnt-like ligands and the postulated Ig-like domains may bind to yet undefined ligand(s), activating distinct second messenger pathways. In addition, the conserved C10 pattern observed in the frizzled-like CRD is predicted to have a distinct three-dimensional structure to the two cysteine-rich (C10) domains of Dickkopf-1 (Glinka et al., 1998) (which also is a Wnt-binding protein), since their cysteine patterns differ.

Conclusion: In conclusion, we have identified a frizzled-like CRD domain within developmental receptor tyrosine kinases, which should help identification of potential ligands and provide insight into the activation of intracellular signaling pathways.

References

- Alcedo J, Ayzenzon M, Von Ohlen T, Noll M, Hooper JE. 1996. The *Drosophila* smoothed gene encodes a seven-pass membrane protein, a putative receptor for the hedgehog signal. *Cell* 86:221-232.
- Altschul SF, Madden TL, Achafer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.
- Bhanot P, Brink M, Samos CH, Hsieh JC, Wang Y, Macke JP, Andrew D, Nathans J, Nusse R. 1996. A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. *Nature* 382:225-230.
- Caruso A, Morris JC, Neben S, Finnerty H, Beier D, Turner K, Wood CR. 1995. Direct submission to GENBANK: <http://www.ncbi.nlm.nih.gov/htbinpost/Entrez/query?form=6&dopt=g&db=p&uid=1017827>.
- Dale TC. 1998. Signal transduction by the Wnt family of ligands. *Biochem J* 329:209-223.
- Ganju P, Walls E, Brennan J, Reith AD. 1995. Cloning and developmental expression of Nsk2, a novel receptor tyrosine kinase implicated in skeletal myogenesis. *Oncogene* 11:281-290.

- Glinka A, Wu W, Delius H, Monaghan AP, Blumenstock C, Niehrs C. 1998. Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* 391:357–362.
- Jennings CG, Dyer SM, Burden SJ. 1993. Muscle-specific trk-related receptor with a kringle domain defines a distinct class of receptor tyrosine kinases. *PNAS* 90: 2895–2899.
- Lewis PM, Crosier KE, Wood CR, Crosier PS. 1996. Analysis of the murine Dtk gene identifies conservation of genomic structure within a new receptor tyrosine kinase subfamily. *Genomics* 31(1):13–19.
- Leyns L, Bouwmeester T, Kim SH, Piccolo S, De Robertis EM. 1997. Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. *Cell* 88:747–756.
- Masiakowski P, Carroll RD. 1992. A novel family of cell surface receptors with tyrosine kinase-like domain. *J Biol Chem* 267:26181–26190.
- Melkonyan HS, Chang WC, Shapiro JP, Mahadevappa M, Fitzpatrick PA, Kiefer MC, Tomei LD, Umansky SR. 1997. SARPs: A family of secreted apoptosis-related proteins. *PNAS* 94(25):13636–13641.
- Oishi I, Sugiyama S, Liu ZJ, Yamamura H, Nishida Y, Minami Y. 1997. A novel Drosophila receptor tyrosine kinase expressed specifically in the nervous system. Unique structural features and implication in developmental signaling. *J Biol Chem* 272:11916–11923.
- Rehn M, Pihlajaniemi T. 1995. Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. The longest form contains a novel motif homologous to rat and Drosophila frizzled proteins. *J Biol Chem* 270:4705–4711.
- Song L, Fricker LD. 1997. Cloning and expression of human carboxypeptidase Z, a novel metalloproteinase. *J Biol Chem* 272:10543–10550.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Valenzuela DM, Stitt TN, DiStefano PS, Rojas E, Mattsson K, Compton DL, Nunez L, Park JS, Stark JL, Gies DR, Thomas S, Le Beau MM, Fernald AA, Copeland NG, Jenkins NA, Burden SJ, Glass DJ, Yancopoulos GD. 1995. Receptor tyrosine kinase specific for the skeletal muscle lineage: Expression in embryonic muscle, at the neuromuscular junction, and after injury. *Neuron* 15:573–584.
- Wang S, Krinks M, Lin K, Luyten FP, Moos M. 1997. Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. *Cell* 88: 757–766.