Chapter 5

Structural Determinants of Binding and Specificity in Transforming Growth Factor-Receptor Interactions.

5.1 Summary

The transforming growth factor (TGFβ) families of proteins are cytokines that occur as large number of homologous proteins. Three major subgroups of these proteins have been realized, the TGFβs, the activins/inhibins and the BMPs, with marked specificities for their receptors. Although structural information is available for some members of TGFβ family of ligands and receptors, very little is known about the way these growth factors interact with the extracellular domains of their cell surface receptors, especially receptor type2. The elements that are determinants of binding and specificity of the ligands are also poorly understood. The structure of the extracellular domain of the receptor is a three-finger fold similar to some of the toxin structures. Amino acid exchanges between multiply aligned homologous sequences of type2 receptors point to residues at the surface, specifically, finger1, as determinant of ligand specificity and complex formation. The 'knuckle' epitope of ligands is predicted as the surface that interacts with the type2 receptor. The residues on strands β2, β3, β7, β8 and the loop region joining β2-β3 and β7-β8 of the ligands are identified as determinants of binding and specificity. These results are independently supported by docking studies of the type2-receptor to the ligand dimer-type1-receptor complex.

5.2 Introduction

5.2.1 The system

The transforming growth factor β (TGFβ) family comprises a large number of structurally related polypeptide growth factors, each capable of regulating a fascinating array of cellular processes including cell proliferation, lineage determination, differentiation, motility, adhesion and death. Expressed in complex temporal and tissue-specific patterns, TGFβ and related factors play a prominent role in development, homeostasis, and repair of virtually all tissues in organisms (Massagué, 1998). For example, the founding member of the TGFβ1 family was identified as a regulator of mesenchymal growth and, separately, as an antimitogen in epithelial cells (Massagué, 1990; Roberts and Sporn, 1993). Activins were identified as endocrine regulators of pituitary function and, as inducers of mesoderm in frogs (Kingsley, 1994; Gaddy-Kurten *et al*., 1995). Bone morphogenic proteins (BMPs) were identified as bone repair factors and, independently, as dorsalizing agents in Drosophila (Hogan, 1996; Mehler *et al*., 1997). Nearly thirty members of the TGFβ family have been described in human and many orthologs are known in mouse, Xenopus and other vertebrates (Massagué, 1998; Hogan, 1996). Four are present in *Caenorhabditis elegans* (Padgett et al., 1998) and seven in *Drosophila melanogaster* (Raftery *et al*., 1999).

The family is divided into two general branches: the BMP/GDF (growth and differentiation factor) and TGFβ/Activin/Nodal branches, whose members have diverse, albeit often complementary effects. Additional members such as inhibin-α act as ligand antagonists. Some family members are expressed in a few cell types or for limited periods of time during development, whereas others are widespread during embryogenesis and in adult tissues. AMH/MIS (Anti-Müllerian hormone or Müllerian inhibiting substance) and GDF8/myostatin are examples of the former; TGFβ1 and BMP4 are of the latter (Massagué *et al*., 2000). TGFβ1-3 are ~70% conserved among themselves, while BMPs are ~60% identical among each other. TGFβs and BMPs share

~30% identity, while inhibinβB shares 30%, 40% and ~27% sequence identity with BMP7, BMP2 and TGFβs, respectively. *Dpp* (*decapentaplegic* protein of *Drosophila melanogaster*) shares 56, 72, 40 and ~31% sequence identity with BMP7, BMP2, inhibinβB and TGFβs, respectively. Glial cell line-derived neurotrophic factor (GDNF) and its subfamily members, undergo similar modes of dimerization as TGFβs, but share very low sequence similarities (~14%) with members of TGFβ family. GDNF subfamily, therefore, can be considered as a member of the broader 'cystine-knot' superfamily, which includes nerve growth factor and platelet-derived growth factor, that have similar protomer structures but display different modes of dimerization and share ~15% sequence similarity with TGFβ family (Saarma, 2000; Sowdhamini *et al*., 1998).

5.2.2 Signal Transduction Pathway

Members of TGFβ family of growth factors are synthesized as larger precursor molecules with an amino-terminal signal sequence and a pro-domain of varying size. These precursor proteins are usually cleaved at a dibasic or RXXR site to release a mature carboxy-terminal segment of 110-140 amino acids (Massagué, 1998; Murray-rust *et al*., 1993; Barr, 1991) and are biologically active as dimers. Members of TGFβ family regulate gene expression by bringing together two types of receptor serine/threonine kinases (Massagué, 1998), collectively known as TGFβ receptor family. Unlike other members of the TGFβ family, GDNF family ligands activate intracellular signaling cascades via the receptor tyrosine kinase Ret (Heldin *et al*., 1997). TGFβ receptor family is divided into two subfamilies: type1 receptors and type2 receptors, on the basis of their structural and functional properties. Table 5.1 summarizes various TGFβ family ligands and their receptors identified biologically. Two general modes of ligand binding have been observed: One mode involves direct ligand binding (The biologically active form of TGFβ ligands is dimer of two monomers. Henceforth, when ligand binding to receptor is discussed, it is assumed that the receptor(s) is interacting with ligand dimer) to ectodomain of the type2 receptor and subsequent interaction of this complex with the type1 receptor. Type1 receptor, in effect, becomes recruited to the complex, which is

characteristic of TGFβ and activin receptors. The second mode of binding is typical of BMP receptors and is cooperative, involving type1 and type2 receptor ectodomains that bind ligand with high affinity when expressed together but with low affinity when expressed separately (Massagué, 1998). In the mechanisms described above, type2 receptors bind to ligand dimers, subsequently (or simultaneously) recruit type1 receptors and finally phosphorylate type1 receptors at GS domain and thus activate them in transducing the signal to the nucleus via SMAD proteins. (Please see reviews by Massagué, 1998; Massagué *et al*., 2000; Miyazono *et al*., 2000; Ducy and Karsenty, 2000; Zimmerman and Padgett, 2000; Massagué, 2000; for the list of TGFβ family members, their activities and detailed signaling mechanism.)

5.2.3 Description of Structure of Ligands

The TGFβ isoforms show remarkable structural homology between each other, including seven absolutely conserved cysteine residues that form three intrachain disulfide bonds and one interchain disulfide bond. TGFβ and activins/inhibins contain an extra disulfide bridge at the N-terminus of the molecule. The structures of TGFβ2 (Daopin *et al*., 1992; Schlunegger and Grutter, 1993), TGFβ3 (Mittal *et al*., 1996), BMP7/OP1 (Griffith *et al*., 1996), BMP2 (Scheufler *et al*., 1999) and GDNF (Eigenbrot and Gerber, 1997) were determined by X-ray crystallography while a model of TGFβ1 (Hinck *et al*., 1996) was calculated from NMR restraints. The monomer is a thin, elongated and slightly curved molecule resembling an open left hand. As shown in Figure5.1a, each monomer is folded into nine β-strands (β1-β9) and a long α-helix (discussed below as α3). The fold can be described as a hand with the thumb as the N-terminus and the extended sheets as fingertips representing β2-β3 and β7-β8 loops. Accordingly, the convex surfaces of the fingers correspond to the knuckles and the helix region to the 'wrist'. The residues exposed on the convex surface involving β2, β3, β7 and β8 strands and loops joining them define the 'knuckle' epitope (Kirsch *et al*., 2000). All known ligand sequences contain seven invariant Cys residues, numbered as C2, C4 and C5-C9; many of them contain an extra pair of Cys residues numbered as C1 and C3. The structurally conserved

region of the fold is described as a '*cystine-knot*' since cysteines C4, C5, C8 and C9 participate in an eight-membered macrocycle wide enough for the last cystine bridge (formed by C3 and C7) to pass through. The cysteines that form N-terminal disulfide bridge in TGFβ are absent in other family members. Since the proteins of this family lack the hydrophobic core, the rigid cystine-knot scaffold is necessary for structural integrity. Further stabilization is achieved by dimerization that creates a hydrophobic core between the protomers. In most cases, such dimerization events are accompanied by the formation of a disulfide bridge connecting the two protomers at the C6 position.

5.2.4 The Structure of Receptors

The type1 and type2 receptors are glycoproteins of approximately 55kDa and 70kDa, respectively, with core polypeptides of 500 to 570 amino acids including the signal sequence. Each receptor contains an extracellular or ectodomain, a short membrane spanning helix and a cytoplasmic serine/threonine kinase domain (Mathews and Vale, 1991; ten Dijke *et al*., 1993; Lin *et al*., 1992; Attisano *et al*., 1992; Ebner *et al*., 1993). The type1 receptors have a higher level of sequence similarity than type2 receptor, particularly in the kinase domain (Massagué, 1998). Crystal structure of the extracellular domain of activin type2A receptor (AtR2-ECD) has been determined (Greenwald *et al*., 1999). The fold of AtR2-ECD comprises of three antiparallel sheets formed by seven β-strands (Figure 5.1b). The molecule has both concave and convex surfaces arising from a curvature in the first β-sheet (β1-β2). AtR2-ECD adopts a *three-finger toxin* fold, also observed in several toxins, which is characterized by a common pattern of eight cysteines, forming a conserved scaffold of four disulfide bridges. The three fingers refer to three pairs of strands (β1-β2, β3-β4, β5-β6) which all point roughly to the same direction (Figure 5.1b). AtR2 and cardiotoxin have the same disulfide pattern (C1-C3, C2-C4, C5-C8 and C9-C10), with the exception of an additional disulfide in AtR2 (C6-C7). Among the type2 receptors, there is some variability in the occurrence of the cysteines. The majority of the extra cysteines in other receptors are clustered in finger1, which constitutes the least conserved region in terms of both sequence and length. TβR2

has four additional cysteines in finger1, but lacks the two cysteines that constitute the C5-C8 disulfide bond in AtR2. Punt has two extra cysteines in finger1.

Crystal structure of human BMP2 ligand in complex with two high affinity receptor1A extracellular domains ($BR1A_{ec}$) has been reported recently which provide important information on TGF-receptor interactions at the molecular level (Kirsch *et al*., 2000). In this structure, two molecules of type1 receptor are bound to the ligand dimer at the 'wrist epitope' region of the ligand (further details please see Results and Discussion) by mainly hydrophobic surfaces of both the molecules. In addition, this report also confirms that both type1 and type2 receptor extracellular domains share the same fold, especially at the central β-sheet, despite poor sequence identity. Differences at loop regions and insertions of non-core secondary structures are evident; for example, a helix involved in primary interactions with the ligand in type1 receptor is absent in type2 receptors (Kirsch *et al*., 2000). Instead, an additional disulfide bridge, unique to type2 receptors, links the equivalent loop region at the convex surface to the central β-sheet. This suggests that the two types of receptors have different modes of binding to the ligand at the atomic level.

5.2.5 Previous Studies and Present Approach

Only a limited number of functionally important residues have been identified in TGFβ and related growth factors for binding to type2 receptor. The influence of segment deletions, residue replacements and isoform chimeras on the binding affinity of TGFβs of their type2 receptor (TβR2) were studied, highlighting the importance of C-terminal residues 83-112 of TGFβ1-3 (Qian *et al*., 1996). Structure-function analysis of activinβA molecule is reported and two amino acids involved in the binding of the activin molecule to its type2 receptor were identified as important for binding: Asp27 and Lys102, on the 'knuckle epitope' (Wuytens *et al*., 1999). Gray and coworkers have performed alanine scanning mutagenesis experiments on AtR2-ECD and identified a cluster of hydrophobic residues ('hydrophobic triad'), Phe42, Trp60 and Phe83, as critical for binding to activins/inhibins (Gray *et al*., 2000). It is known that type2 receptors form a heteromeric complex with the ligand, but exactly how many receptor molecules interact with the

ligand is not known (Massagué, 1998). It is apparent from Table 5.1 that TGFβ ligands can only bind to TβR1 and TβR2 but no such specificity is observed in the case of BMPs and activins. AtR1 binds to activins/inhibins, BMP7 and MIS/AMH; AtR2 binds to activins, BMP7 and GDF5 (Massagué, 1998). This report suggests that the determinants of ligand binding to receptors may be conserved within the TGFβ subfamily, the determinants of specificity are different between TGFβ and activin/BMP subfamilies, while activins and BMPs have similar residues that determine the specificity. Activin receptors bind to activins/inhibins, BMPs, MIS and GDF5. Despite a remarkable structural similarity, no such binding is observed for TGFβ ligands.

In order to determine the functionally important residues*,* we have compared the sequence distribution within the three fingers of the receptor, the nature of charge distribution of ligands and employ the Evolutionary Trace (ET), first applied by Litcharge *et al* on SH2 and SH3 domains method (Litcharge *et al*., 1996), to identify potential binding-site residues as targets for mutagenesis in TGFβ family of receptors. The five available structures of TGFβ ligands (3TGFβs, and 2BMPs) and three-dimensional models of *dpp* and inhibinβB derived by comparative modeling, have been analyzed for the differences in the distribution of polar and hydrophobic residues on the surface of the molecules, especially at the conserved residues (Innis et al., 2000) important for binding to type2 receptor. The extracellular domain of type2-receptor was docked to the ligand-dimer-type1 receptor complex. On the basis of previous mutagenesis studies and the results of our analysis, the 'knuckle' epitope is identified as a site of interaction with type2 receptor. Since the ligand molecules contain two symmetric knuckle epitopes, two receptors can bind to ligand dimer forming a tetrameric complex.

5.3 Materials and Methods:

5.3.1 Sequence alignment and clustering of receptor2 sequences

23 members of TGFβ receptor2 family were identified by PSIBLAST (Altschul *et al*., 1997) search using AtR2A ectodomain as query sequence against the Swissprot Databank (Bairoch and Apweiler, 1996) and used for evolutionary analysis. The ectodomains of the sequences were aligned using CLUSTALX (V 1.8; Thompson *et al*., 1997) and manually edited ensuring that gaps were not inserted into areas of known (or predicted) secondary structures. A PHYLIP (V 3.5) distance matrix based on sequence dissimilarity indices was generated and input into KITSCH clustering package to build a rooted phylogenetic tree (Felsenstein, 1985).

5.3.2 Evolutionary Trace analysis of receptor sequences

An evolutionary trace is generated by comparing consensus sequences for a group of proteins which originate from a common node in a phylogenetic tree and are characterized by a common Evolutionary Time Cut-off (ETC), and classifying each residue as one of the three types: absolutely conserved, class-specific and neutral. Here 'class-specific' denotes residues occupying a strictly conserved location in the sequence alignment, but differing in the nature of their conservation between various subgroups. When structural and functional residues of a protein family are not characterized, target residues can be chosen for mutagenesis. This can also be mapped on to known protein structures to identify clusters of important amino acids on the surface of the protein.

The ET analysis (Litcharge *et al*., 1996) was performed using TraceSuite (Innis et al., 2000). First, the phylogenetic tree was split along the evolutionary time into five evenly distributed partitions: P01 to P05 in order of increasing ETC. For each partition, a trace procedure was completed automatically in three steps: (1) Protein connected by a common node with evolutionary time greater than the given ETC were clustered together.

(2) A consensus sequence was generated for each group to distinguish between conserved and non-conserved positions. (3) A trace was generated by comparing the consensus sequences of receptors. Residues were classified into three types: absolutely conserved, class-specific and neutral. All the receptor sequences considered for the initial alignment were used for ET analysis. Punt sequence was not included since it is a lone element in the evolutionary tree and may bias the results.

5.3.3 Comparative modeling and visualization

Mature carboxy terminus peptides of *dpp* of *Drosophila melanogaster* and inhibinβB of *Homo sapiens* were taken from Swissprot databank (Bairoch and Apweiler, 1996). They were multiply aligned using CLUSTALX (V 1.8; Thompson et al., 1997) to other family members of TGFβ family. MODELLER (V 4.0; Sali and Blundell, 1993) was used to build three-dimensional models of both the proteins. BMP2 (PDB code 3bmp) was used as a template for modelling *dpp*; BMP2 (PDB code 3bmp) and TGFβ3 (PDB code 1tgj) were used as templates for modeling inhibinβB. MODELLER constructs a minimized 3D model(s) of a protein by the satisfaction of spatial restraints extracted from the template PDB (Bernstein *et al*., 1977) files. 20 models of the query sequence in each case were generated. The final models were chosen on the basis of lowest energy and least violation of structural restraints. The models with violated backbone CO and backbone NH restraints are not considered. Stereochemistry and geometry of the models were assessed using PROCHECK (V 3.4.4; Laskowaski *et al*., 1993) ensuring that the models have more than 85% residues in the core region of Ramachandran plot. The models were energy minimized using MAXIMIN2 option in SYBYL (Tripos Association, Inc., V6.5) using TRIPOS force field. For every run of energy minimization, 20 cycles of Simplex method and a further 50 cycles of Powell algorithm were employed. The resultant models have no short contacts or bad geometry. The dimer coordinates were generated using a superposition program called SUPER (Neela, B., personal communication). The punt (type2 receptor for *dpp* molecule) receptor ectodomain was also modeled following the same procedure with AtR2-ECD crystal structure (PDB code 1bte) as template. The resultant models and crystal structures were viewed by RASMOL (V 2.6b2; Sayle and

Milner-white, 1995) and solvent accessible surfaces and electrostatic potentials were calculated and displayed using GRASP (V 1.1; Nicholls *et al*., 1993). Structure-based sequence alignment of TGFβ ligands was compiled using the program COMPARER (V 2.0; Sali and Blundell, 1990) and structure-annotated using JOY (V 4.0; Overington *et al*., 1993; Mizuguchi *et al*., 1998).

5.3.4 Docking studies on ligand-receptor type2 receptor interactions

The Global Range Molecular Matching (GRAMM, V 1.03) methodology (Katchalski-Katzir *et al*., 1992; Vakser, 1995; Vakser, 1996) is an empirical approach to smoothing the intermolecular energy function by changing the range of the atom-atom potentials. The technique allows to locate the area of the global minimum of intermolecular energy for structures of different accuracy. The quality of the prediction depends on the accuracy of the structures. Thus, the docking of high-resolution structures with small conformational changes yields an accurate prediction, while the docking of ultra-low-resolution structures will give only the gross features of the complex. To predict the structure of a complex, it requires only the atomic coordinates of the two molecules (no information about the binding sites is needed). The program performs an exhaustive 6-dimensional search through the relative translations and rotations of the molecules.

The X-ray structures of activin type2 receptor (PDB code 1bte; solved at 1.5 Å resolution) and complex of BMP ligand dimer with its type1 receptors (PDB code 1es7; solved at 2.90 Å resolution) were docked using GRAMM program (Katchalski-Katzir *et al*., 1992; Vakser, 1995; Vakser, 1996) with a generic, hydrophobic mode and a grid step of 2.1Å. 1000 different models were generated to study every probable way of ligand-receptor interactions. The models were examined for maximal hydrophobic interactions and total interactions between the 1es7 and 1bte structures using the distance cut-off value derived from known cytokine-receptor crystal structures.

5.4 Results and Discussion

5.4.1 Analysis of receptor type2 sequences

Aligned non-redundant sequences of receptor type2, as shown in Figure 5.2, contain seven sequences of AtR2A, seven sequences of AtR2B, four BR2 sequences, four TβR2 sequences, and a punt receptor sequence from *Drosophila melanogaster*. Sequences of subfamilies show high conservation among themselves, but across the subfamily there is hardly any conservation apart from the cysteines. Phe42, Trp45, Gly58 and Asn92 are characteristic of a three-finger toxin fold and are largely conserved. However, Phe42 is substituted by Tyr in AtR2B, BR2 and punt but replaced by Val in TβR2; Gly58, which is conserved in AtR2B, BR2 and punt, is absent in TβR2. In general, the average sequence identity is around 25%. Punt receptor shares 28-30% identity with BR2 and AtR2B, 22-23% identity with AtR2A and ~15% identity with TβR2. Trp45 and Asn92 are absolutely conserved amongst all the type2 receptor subtypes considered. Evolutionary tree was generated using PHYLIP3.5 package (Figure 5.3; Felenstein, 1985). As expected, AtR and BR sequences are more similar and TβRs stand by their own as a separate cluster.

5.4.2 Analysis of residues in fingers

Finger1 contains loops of similar length that may be important for specificity in binding to the ligand since these loop regions display maximal sequence variation also confirmed by evolutionary trace method (discussed later). Two negatively charged residues at the tip of finger1, Glu19 and Asp21 (AtR2A numbering) are replaced by Asn and Leu in BR2 and Ser and Cys in TβR2. Finger1 of punt contains an extra disulfide bridge, while that of TβR2 contains two extra disulfide bridges (see Figure 5.1). This confirms previous modeling and scanning-deletion mutagenesis studies Guimond *et al*., 1999), which show that residues in finger1 (residues 58-60 and 63-65 of TβR2), facing the concave surface are important to bind TGF. Finger2 contains very few residues in each receptor sequence.

However, punt, TβR2 and BR2 receptors have relatively longer finger2 region: two-residue insertion in the case of TβR2 and punt and a one-residue insertion in BR2. Residues 74-79 of finger3 are exposed on the concave surface; AtR2A has two positive and two negative charges in this loop, while AtR2B is predominately negative. BR2 is polar and TβR2 is predominantly positive in this region, while punt contains one positive and one negatively charged residue.

It is reported that mutant receptors, containing deletions corresponding to loop regions of finger1, β2-β3 loop and finger2, do not bind the ligand (Guimond *et al*., 1999). However, mutant receptors containing deletion at finger3, loop region before β1, β4-β5 loop and after β7 do bind the ligand with similar affinities as the wild type receptors (Guimond *et al*., 1999). Deletion of the loop region corresponding to finger2, owing to the fact that finger2 is short, might cause structural changes to the receptor rendering inability to bind the ligand. Thus, finger2 may or may not be important for binding. The highly variable finger1 is not only a potential binding interface, but also the second most exposed, conserved hydrophobic surface (as observed in the crystal structure of AtR2), which is present at the convex side of the molecule. Finger1 is a good candidate to provide both hydrophobic docking surface and to act as primary determinants of interaction and binding specificity (Greenwald *et al*., 1999).

5.4.3 Evolutionary Trace of receptor2 sequences

The output of TRACESUITE program (by Innis *et al;* employing ET method; Litcharge *et al*., 1996) on the extracellular domain of TGF type2 receptors is shown in Figure 5.4. Analysis of the mapped traces for partitions P01 to P05 reveal clusters of potentially important residues on both concave and convex surfaces of the receptor structures. The residues defined by the 'hydrophobic triad' are located at the concave surface (Gray *et al*., 2000). In partition P01, apart from the structurally invariant cysteines, Trp45, Val55 and Asn92 are absolutely conserved among all receptor types considered in the ET analysis. The conserved Val55 lies on β4 and it is in the vicinity of finger2. Lys replaces Val55 in punt sequence. Other residues identified in partition P01 are Thr8, Glu10, Asn15, Glu19,

Glu29, Gly33, Ala43, Asn47, Asp62, Asp63, Val81, Glu93 and Phe95. Gly33 and Ala43 are buried in the core and may have a structural role and Glu19 is on the β1-β2 loop (finger1). Thr8, Glu10, Asn15, Asn47, Glu93 and Phe95 do not face the concave surface but are solvent accessible with no identified function. The trace residues facing the concave surface are Glu29 (on β2), Asp62 and Asp63 (on β4-β5 loop), Val81 (on β5-β6 loop; finger3) and Phe83 (on β6; finger3). No class-specific residues were identified at P02.

Phe83 is in the 'hydrophobic triad' identified by alanine scanning mutagenesis to be important for ligand binding (Gray *et al*., 2000). However, single mutations of Phe13, Phe14, Glu29 and Asp62 do not alter binding specificity for activins and inhibins (Gray *et al*., 2000). ET method does not identify Phe13 (exposed on concave side) and Phe14 (exposed in convex side) as trace residues, which implies that these residues are probably involved in non-specific binding. The method, however, identifies Glu29 (at the end of β1) and Asp62 (β4-β5 loop), which face away from the three fingers (Figure 5.1b). Glu29 is replaced by Ser in both TβR2 and BR2 and by Thr in punt; Asp62 is replaced by Gly in BR2 and by Tyr in TβR2 while the corresponding residue in punt is deleted. This suggests that Glu29 and Asp62 might be playing a functional role in other subfamilies not tested so far by mutagenesis experiments.

In the crystal structure of AtR2A (Greenwald *et al*., 1999), Thr44 (AtR2A numbering) identified as a conserved residue at partition P03 by ET analysis, is in the middle of a solvent-exposed hydrophobic surface, created by Ala16, Phe42, Val55, Trp60, Ile64, Val81 and Phe83 (to recall that three of these define the 'hydrophobic triad' important in ligand binding). Except Ala16 all others are 'trace' residues. Ala16 is considered in the analysis as it is solvent exposed hydrophobic residue and it lies in loop region of finger1. We refer to the 'hydrophobic triad' as the residues defining the 'principal' hydrophobic patch which can be further extended to include Ala16, Thr44, Val55, Leu61, Ile64 and Val81 termed as the 'surrounding' hydrophobic patch. In BR2 and punt, Thr44 is replaced by Leu in BR2 and a Val in TβR2. To note that position Lys56, spatially proximate to this extended hydrophobic patch and conserved in AtR, BR and punt when mutated to

Ala does not display drastic change in binding (Gray *et al*., 2000). Lys56 has not been identified as a trace residue by our present ET analysis.

5.4.4 Structure based analysis of TGFβ ligands and identification of determinants of binding and specificity

Large exposed hydrophobic patches on a protein surface often form part of a binding surface (Young *et al*., 1994). In the human growth hormone-receptor complex, a few hydrophobic residues at the interface contribute most to the free energy of interaction (Clackson and Wells, 1995). The recently solved crystal structure of the complex of BMP2-BR1A ectodomain (Kirsch *et al*., 2000), exemplifying TGF-TGF type1 receptor interactions, also demonstrates the same theme. Phe85 of $BR1A_{ec}$ helix α1 fits into a hydrophobic pocket of the ligand where it interacts with Trp28 and Trp31 of BMP2, among other residues. In the crystal structure of free BMP2, this pocket accommodates a 2-methylpentane-2, 4-diol molecule from the buffer solution, and a dioxane in the case of TGFβ3 (Mittal *et al*., 1996; Scheufler *et al*., 1999; Kirsch *et al*., 2000). Ile62, Val63, and Leu66 of BMP2 provide an almost exclusively hydrophobic surface, which together with Asn59, form the site of interaction with Phe85 of the receptor molecule (Kirsch *et al*., 2000). In addition, Phe60, Met78 and Ile99 of BR1A are central to the ligand-binding interface (Kirsch *et al*., 2000). The residues correspond to Asn59, Ile62, Val63, and Leu66 (BMP2) in case of TGFβ ligands (Innis *et al*., 2000) and the residues corresponds to Phe85, Phe60, Met78 and Ile99 (BR1A) in case of receptor1 sequences were identified as trace residues. In order to identify the determinants of binding and specificity for TGF-TGF type2 receptors, the following approaches were taken:

5.4.5 Structure of TGF growth factors and analysis of TGF-like sequences

The structures for TGFβ1-β3 (Daopin *et al*., 1992; Schlunegger *et al*., 1993; Mittal *et al*., 1996; Hinck *et al*., 1996), BMP7 (Griffith *et al*., 1996) and BMP2 (Scheufler *et al*., 1999) when superposed in the best fit, display an overall root mean square deviation of less than

1.1 A°. However, there are clear differences in some structural elements between TGFβs and BMPs; N-terminus is not visible in the crystal structure of BMP2 (Scheufler *et al*., 1999) and BMP7 (Griffith *et al*., 1996). In contrast, TGFβ1-3 exhibits a short N-terminal α-helix (α1), that is anchored to the protein core by an additional disulfide bridge (Daopin *et al*., 1992; Schlunegger *et al*., 1993; Mittal *et al*., 1996; Hinck *et al*., 1996). Moreover, BMP2 and BMP7 do not contain the short helix α2 observed after the second β-strand in TGFβs and is replaced by a tighter non-helical turn. This feature is conserved among known BMPs, GDFs, activins and other subfamilies. However, BMP2 and BMP7 structures show a unique conformation at the loop preceding α3: a longer loop with a three-residue insertion (a short β-strand in BMP2).

5.4.6 Analysis of surface residues of ligand molecules for difference in charge distribution

Figure 5.5 shows the GRASP surface representation (Nicholls *et al*., 1993) of structures of TGFβ1-3, BMP2, BMP7 and models of inhibinβB and *dpp*. Large hydrophobic areas are concentrated especially on the wrist and knuckle epitope regions of the ligand dimers. It is clear from the figure that the charge distribution is different between TGFβ isoforms and activin subfamily of proteins, especially in the knuckle epitope, in the loops of β2-β3 and β7-β8 strands. These regions contain high negative charge in case of BMPs, inhibins and *Dpp*, while they are positively charged in TGFβs (the conservation is confirmed using multiple sequence alignment of ligands). Unlike BMPs, *dpp* is in general polar at β7-β8 loop, at the knuckle epitope, where Asp93, Glu95 and Lys96 of BMP2 are replaced by Asn, Gln and Thr in *dpp*, respectively. This difference in charge distribution together with the structural differences discussed before can be instrumental in giving rise to specificity while binding to receptors. In addition, all the structures have positive charge at N-terminus (conserved positively charged residue, two residues after C2) which accounts for their heparin binding (Ruppert et al., 1996). However in BMPs, the N-terminus might fold back to shield this charge (as observed in case of inhibinβB model) and TGFβs are less positive than BMPs in this region.

5.4.7 Evolutionary Trace of Ligands and identification of residues implicated in binding and specificity:

ET method was applied to multiply aligned sequences of TGFβ superfamily of ligands and trace residues were identified by Innis and coworkers (Innis *et al*., 2000). Trp28 and Trp31 of BMP2, which have primary interactions with Phe85 of BR1Aec (please see above) are absolutely conserved in the TGFβ family alignment and are identified as 'trace' residues (Innis *et al*., 2000). Mutation of Trp31 to alanine significantly decreases the stability of the BMP2-BR1A_{ec} complex (Kirsch *et al*., 2000). Interestingly, neither Trp28 nor Trp31 are conserved in the distant relatives like GDNF.

Here we will discuss those trace residues which occur on the knuckle epitope (important for receptor binding), are topologically equivalent (using COMPARER; Sali and Blundell, 1990) and display similar characteristics (identified by JOY; Overington et al., 1993; Mizuguchi et al., 1998) (Figure 5.6). Two interesting clusters of residues are identified: first at alignment positions 35, 36, 37, 92, 93, 94, 96, 98, 104, 105 and 106 (Figure 5.6; residues forming the 'knuckle epitope'). In case of BMP2 (PDB code 3bmp), these residues are Val33, Ala34, Pro35, Ala86, Ile87, Ser88, Leu90, Leu92, Val98, Val99 and Leu100. A second small cluster of residues at alignment positions 17, 18, 43 and 115 (Arg16, His17, Phe41 and Glu109 according to BMP2 numbering) is rather surprising. However, it should be noted that the N-terminus of the BMP family of ligands, that are not seen in the crystal structure could fold back in this region attributing a structural than functional role to these residues. The trace residues at the 'knuckle epitope' are divided into two classes (see Table 5.3 a;b): First, the residues, which are implicated for ligand binding and second, residues implicated for binding as well as specificity, most of which are subfamily specific and class-specific residues. The first cluster can be the preferred site of interaction with hydrophobic clusters on type2 receptor identified by mutagenesis studies and also in this study. Residue Pro35 is absolutely conserved in all the ligand sequences considered for our analysis but is absent in GDNF. Val33 and Ala34 are replaced by charged residues in TGFβ isoforms, while Leu100, is replaced by a charged residue in inhibins; other residues in cluster1 are conserved substitutions. In cluster2,

Arg16 is replaced by a hydrophobic residue in TGFβs. Alignment position 18 (Figure 5.6; His17 of BMP2) is occupied by a positively charged residue in all the known sequences. Phe41, adopting unusual φ/ψ angles (67°, 178°) in the Ramachandran plot (Ramachandran *et al*., 1963; Ramachandran and Sasiekharan, 1968), although not in the dimer interface, forms interchain contact at the backbone carbonyl with the neighboring subunit in bmp2 and its side chain is solvent exposed (Scheufler *et al*., 1999) on 'knuckle epitope'. This residue is present where the β-strands, β2, β5, β6 and β9 are arranged close enough to form a short segment of four-stranded antiparallel β-sheet as evident in the crystal structure (Scheufler *et al*., 1999). This arrangement is also observed in all known proteins in TGFβ superfamily. As shown in Table 5.3, Glu109 (BMP2 numbering) is specific for BMP2 and inhibin subfamilies, which is a positively charged residue (Arg/Lys) in TGFβs, BMP7 and a valine in *dpp*. Such differences at 'trace' residues point to class-specific electrostatic distribution and receptor specificity. Figure 5.7 shows the 'trace residues' mapped on the surface of ligand structure. The residues mapped are on 'knuckle epitope'.

In the light of the above results and the crystal structure of BMP2 with its type1 receptor, it is plausible to propose that finger1 (with loop region of β2-β3) of type2 receptor interacts with the 'knuckle epitope' of the ligand; if C-terminus of both type1 and type2 receptors need to point roughly in the same direction and the type2 receptor interacts with the ligand at its concave surface (Greenwald *et al*., 1999; Kirsch *et al*., 2000; Guimond *et al*., 1999) with the 'hydrophobic triad' residues. Finger3 loop region of the type2 receptor is not playing any important role in binding. Some of residues identified by ET method in 'surrounding patch' for receptor type2 are also reported in the deletion studies of core region (53-55, 83-85, 98-100 and 143-145; TβR2 numbering; See table 5.2) of TβR2 (Guimond *et al*., 1999). Thus ET method can be used both to rationalize the result of the mutagenesis studies and also to predict the targets for the mutagenesis.

5.4.8 Docking studies of type2-receptor to the ligand-dimer-type1 complex

The nature of interactions at the ligand-type1 receptor binding site were primarily hydrophobic (Kirsch *et al*., 2000). The above analyses on trace residues of solvent-exposed hydrophobics and the overall similarities between type1 and type2 receptors suggest similar hydrophobic interactions in TGF ligand-type2 receptor binding (Figure 5.7a,b). 1000 GRAMM (Katchalski-Katzir *et al*., 1992; Vakser, 1995; Vakser, 1996) models of type2 receptor interacting with the ligand dimer complexed with two type1 receptors were generated. The models were examined for maximal hydrophobic interactions, (defined as interaction between hydrophobic residues of both the structures) and total interactions (defined as interaction between all residues of both the structures) using the C_α distance cutoff of 12 Å for 'interacting' pairs. Figure 5.7c shows the distribution of the number of models with different number of hydrophobic interactions between ligand-type1 receptor complex and receptor type2. Models with appreciable number of hydrophobic interactions at the 'principal patch' were specifically examined after including the 'surrounding hydrophobic patch' residues (inset to Figure 5.7c). Interestingly, models with the highest number of hydrophobic interactions (Figure8c) and total interactions (as shown in Figure5.7d) with key residues of type2 receptor closely correspond to ET-results, suggesting a theme involving knuckle epitope at the ligand as the receptor-binding site.

5.5 Conclusions

In this chapter, the clusters of residues has been identified, which lie on the knuckle epitope of ligand molecules and the concave surface of type2 receptor molecules, which may play an important role in complex formation. These clusters are hydrophobic patches surrounded by charged residues on the surface of molecules. Finger1 and a part of finger2 of the type 2 receptor, with the central hydrophobic patch, interact with the 'knuckle epitope' of the ligand (mainly convex side on β2-β3, β7-β8 and the loop regions joining them) as it provides the large conserved hydrophobic surface for docking. The β2-β3

loop region may be interacting with the smaller cluster identified by ET bearing good agreement with GRAMM docking studies. While each type1 receptor interacts simultaneously with both the ligand protomers at the 'wrist' epitope (Kirsch *et al*., 2000), we predict that the type2 receptor interactions are with one protomer each at the 'knuckle' epitope. These predictions are supported by deletion studies on ligands (Hinck *et al*., 1996; Qian *et al*., 1996; Gray *et al*., 2000), deletion and mutagenesis studies on receptor type2 sequences (Gray *et al*., 2000; Guimond *et al*., 1999) and orientation of receptor type1 in the crystal structure in complex with BMP2 (Kirsch *et al*., 2000). The amino acids that emerge from the ET method as important for function can be targets for future mutagenesis studies. It will also be interesting to prepare TGF chimeras of loop region between β2-β3 loop and β7-β8 loop since difference in charge distribution in this region may contribute to specificity in identification of receptors. Various tools such as the study of evolutionary trees, conserved residues of the aligned sequences, spatial positions of interesting residues, charge distribution on their three-dimensional fold and docking studies have been employed to provide structural explanations for ligand-receptor specificity which have general value in the area of protein-protein interactions.

Table5.1 Abbreviations include: TGFβ (transforming growth factor β), BMP (bone morphogenic protein), Dpp (decapentaplegic), GDF (growth and differentiation factor), MIS/AMH (Mullerian inhibiting substance/anti mullerian hormone) TβR (transforming growth factor receptor) AtR (activin receptor) and BR (bone morphogenic protein receptor). This table was compiled according to Massague [1], with a few corrections. It should be noted that there is no species specificity observed in ligand-receptor interaction.

Sequential Binding		
Ligand	**Type 2 receptor**	**Type1 receptor**
TGFβ	TβR2	ALK1, ALK2?, TβR1, ALK7
Activins	AtR2, AtR2B	AtR1 , AtR1B
BMP7	AtR2, AtR2B	AtR1
GDF5	AtR2, AtR2B	AtR1 , AtR1B
MIS/AMH	AMHR	AtR1?
Co-operative Binding		
Ligand	**Type 2 receptor**	**Type1 receptor**
BMPs	BR2	BR1A ,BR1B
Dpp	Punt	Thick veins (tkv) Saxophone (sax)
GDF5		AtR1, BR1B

Table 5.2: Residues of receptor type2 important for interaction with ligands identified using ET method.

	Prin. Patch			Surrounding patch					
Position	91	111	146	58	93	106	116	119	139
AtR2A	F42	W60	F83	A16	T44	V55	L61	I64	V81
AtR2B	Y42	W60	F84	A16	S44	V55	L61	F64	V81
BR2	Y41	W59	F89	P13	L43	V54	I62	P65	I82
TβR2	V85	H102	M135	V56	V87	V100	T108	G110	G130
Punt	Y37	F57	F81	E10	L39	K52	T58	M60	G77

Table 5.3:

a) Specific residues implicated for high affinity binding

Position	21	23	25	37	43	94	96	98
Tgfβ3	Y21	D23	R25	P36	N42	T87	L89	Y91
Tgfβ2	Y21	D23	K25	P36	N42	T87	L89	Y91
Tgfβ1	Y21	D23	R25	P36	N37	P87	V89	Y91
BMP2	Y20	D22	S24	P35	F41	S88	L90	L92
BMP7	Y44	S46	R48	P59	Y65	S113	L115	F117
Dpp	Y9	D11	S13	P24	Y30	A78	L80	L82
IHB β	F17	D19	R21	P32	N39	S89	L91	F93

b) Subfamily specific residues implicated for receptor specificity and binding

Position	17	18	32	35	36	92	93	104
Tgfβ3	V17	R18	K31	H34	E35	P85	L86	P96
Tgfβ2	L17	R18	K31	H34	E35	P85	L86	P96
Tgfβ1	V17	R18	K31	H34	E35	P85	L86	P96
BMP2	R16	H17	D30	V33	A34	A86	I87	V98
BMP7	K40	H41	D54	I57	A58	A111	I112	V123
Dpp	R5	H6	D19	V22	A23	S76	V77	V88
IHB β	R13	Q14	D27	I30	A31	T87	M88	I99

Position	105	106	107	108	115
Tgfβ3	K97	V98	E99	Q100	K107
Tgfβ2	K97	I98	E99	Q100	K107
Tgfβ1	K97	V98	E99	Q100	R107
BMP2	V99	L100	K101	N102	E109
BMP7	I124	L125	K126	K127	R134
Dpp	V89	L90	K91	N92	V99
IHB β	V100	K101	R102	D103	E110

Figure Legends

Figure 5.1: Ribbon representation of (a) BMP2 ligand (3bmp.pdb) and (b) AtR2-ECD (PDB code: 1bte). Secondary structures are labeled. The knuckle and wrist epitopes are marked on TGF; fingers are marked on the receptor structure. Figure is prepared using MOLSCRIPT (Kraulis, 1991).

Figure 5.2: Multiple alignment of 23 sequences of various type2 receptor ecto-domain with secondary structure and fingers marked.

ActR2A_MOUSE*, ActR2A_RAT, ActR2A_HUMAN, ActR2A_BOVIN, ActR2A_SHEEP, ActR2A_GALLUS, ActR2A_XENLA: activin receptor type2A from mouse, rat, human, bovin, sheep, chicken and xenopus.

ActR2B_MOUSE, ActR2B_RAT, ActR2B_HUMAN, ActR2B_BOVIN, ActR2B_GALLUS, ActR2B_ZEBRAFISH, ActR2B_GOLDFISH: activin receptor type2B from mouse, rat, human, bovin, chicken, zebrafish and goldfish.

BMPR2_HUMAN, BMPR2_MOUSE, BMPR2_GALLUS, BMPR2_XENLA: bone morphogenic protein receptor type2 from human, mouse, chicken and xenopus.

TGR2_HUMAN, TGR2_PIG, TGR2_MOUSE, TGR2_RAT: transforming growth factor receptor type2 from human, pig, mouse, rat and fruitfly.

PUNT_DROSO: homologue of activin type2 receptor in fruit fly.

* ActR2A_MOUSE (PDB code 1bte) is activin type2 receptor ecto-domain sequence from mouse, with known crystal structure (Greenwald *et al*., 1999),. The sequence is showed in structure based annotation using JOY (Overington *et al*., 1993; Mizuguchi *et al*., 1998). Please refer to legend of Figure6 for the JOY key.

Figure 5.3: Dendrogram containing 23 TGFβ family of receptor type2 ecto-domain on the basis of their sequence dissimilarity using PHYLIP3.5 .(Felsenstein, J, 1985) The sequences are as described in legend of Figure2.

Figure 5.4: Evolutionary Trace of type2 receptor sequences (excluding punt) for partitions P01 to P05, aligned with the amino acid sequences of AtR2A_mouse (activin receptor2A from mouse), AtR2B_human (activin receptor2B from human), BR2_human (BMP receptor2 from human) and TβR2 (TGFβ type2 receptor from mouse). * indicates

the residues important for mutagenesis. n indicates solvent buried residues as shown in the crystal structure of AtR2-ECD.(Greenwald *et al*., 1999).

Figure 5.5: Electrostatic potential representation of the known and modeled structures of TGFβ family of ligands using GRASP (Nicholls, 1993). Acidic residues are indicated by red surface patches and blue patches indicate basic residues. The structures are (a) bone morphogenic protein 2, (b) inhibinβB, (c) bone morphogenic protein 7, (d) decapentaplegic protein (e) transforming growth factor β1, (f) transforming growth factor β2 and (g) transforming growth factor β3. The charged residues are marked and numberings are according to their structural positions given in PDB files. In case of *dpp* the hydrophobic residues are marked.

Figure 5.6: Structure based sequence alignment of the TGFβ family using COMPARER (Sali and Blundell, 1990) and compiled using JOY (Overington *et al*., 1990; Mizuguchi *et al*., 1998). Solvent-accessible and solvent-inaccessible residues are shown in upper case and lower case, respectively. Residues in positive phi are indicated in italics; residues with *cis* peptide in the backbone or disulfide bonds are indicated by the presence of breve (e.g. š) or cedilla (e.g. ç), respectively. Hydrogen bonds formed to the side chains, main chain amides and main chain carbonyls of the other residues are indicated by the presence of tilde on top, boldface or underline respectively. The secondary structures are marked and numbered.

Figure 5.7: Predicted mode of interaction between tranforming growth factor and type2 receptor.

a) Structure of bmp dimer in GRASP surface representation (Nicholls, 1993). Key residues identified by evolutionary trace (ET) method (Litcharge *et al*., 1996), probably important for binding (prinicipal patch) are denoted in cyan. Additional residues (surrounding patch), also identified by ET method are shown in magenta.

b) Same as (a) but for the extracellular domain of activin type2 receptor. Regions corresponding to finger1, finger2 and finger3 (also see Figure 1) are shown in yellow, green and violet arrows. Proposed complimentary areas of interaction in the ligand dimer

are marked in similar colors in (a).

c) Distribution of the total number of hydrophobic contacts for 1000 models of the interaction between type2 receptor (PDB code 1bte) and ligand-dimer complexed with two type1 receptor molecules (PDB code 1es7). Hydrophobic contacts are measured between the two molecules as the number of C^{α}-C^{α} distances of hydrophobic residues of 1es7 within 12Å from key residues on the type2 receptor. Key residues at the type2 receptor have been identified by ET method (Litcharge *et al*., 1996; also listed in Table 5.2) and also by alanine-scanning mutagenesis (Gray *et al*., 2000) to be important for binding. Inset to Figure 7c) Models with high hydrophobic contacts (7 or more) at the 'principal patch' of the type2 receptor are examined for additional hydrophobic interactions including the 'surrounding hydrophobic patch'.

d) Ribbon representation of one of the models of the interaction between type2-receptor and type1-receptor-bound ligand dimer. This model, suggested by GRAMM, has high number of hydrophobic residues at the predicted binding site (Figure 7c). Activin type2 receptors are shown in cyan, bmp dimer in magenta and the two type1 receptors are shown in grey. The knuckle epitope of the ligand dimer and finger1 and finger3 of the type2-receptor (shown in yellow and violet arrows) are the key regions predicted to form the binding interface. This picture has been prepared using SETOR (Evans, 1993).

5.6 References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res*. **25**: 3389-3402.

Attisano, L., Wrana, J.L., Cheifetz, S., and Massague, J. (1992). Novel activin receptors: distinct genes and alternative mRNA splicing generate a repertoire of serine/threonine kinase receptors. *Cell*. **10** 68:97-108.

Bairoch, A. and Apwiler, R. (1996). The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucl. Acids Res*. **24**: 21-25.

Barr, P.J. (1991). Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell*. **66**:1-3.

Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.F. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol*. **112**: 535-542.

Clackson, T. and Wells, J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science* **267**: 383-386.

Daopin, S., Piez, K.A., Ogawa, Y. and Daevis, D.R. (1992). Crystal structure of transforming growth factor-β2: an unusual fold for the superfamily. *Science* **257**: 369-374.

Ducy, P., and Karsenty, G. (2000). The family of bone morphogenetic proteins. *Kidney Int* **57**: 2207-14.

Ebner, R., Chen, R.H., Lawler, S., Zioncheck, T., and Derynck, R. (1993). Determination of type I receptor specificity by the type II receptors for TGF-β or activin. *Science*. **262**: 900-2.

Eigenbrot, C., and Gerber, N. (1997). X-ray structure of glial cell-derived neurotrophic factor at 1.9 A resolution and implications for receptor binding. *Nat. Struct. Biol.* **4**: 435-8.

Evans, S. V. (1993). SETOR--hardware-lighted 3-dimensional solid model representation of macromolecules. *J. Mol. Graph*. **11**: 134-138.

Felsenstein, J. (1985). Confidence-limits on phylogenies - An approach using the bootstrap. *Evolution* **39**: 783-791.

Gaddy-Kurten, D., Tsuchida, K, and Vale W. (1995). Activins and the receptor serine kinase superfamily. *Recent Prog Horm Res* **50**:109-29.

Gray, P.C., Greenwald, J., Blount, A.L., Kunitake, K.S., Donaldson, C.J., Choe, S. and Vale, W. (2000). Identification of a binding site on the type II activin receptor for activin and inhibin. *J.Biol.Chem*. **275**: 3206-3212.

Greenwald, J., Fischer, W.H., Vale, W.W. and Choe S. (1999). Three-finger toxin fold for the extracellular ligand-binding domain of the type II activin receptor serine kinase. *Nat*. *Struct. Biol*. **6**: 18-22.

Griffith, D.L., Keck, P.C., Sampath, T.K., Rueger, D.C. and Carlson, W.D. (1996). Three-dimensional structure of recombinant human osteogenic protein 1: structural paradigm for the transforming growth factor β superfamily. *Proc. Natl. Acad. Sci*. USA **93**: 878-883.

Guimond, A., Sulea, T., Pepin, M.C.O. and Connor-McCourt, M.D. (1999). Mapping of putative binding sites on the ectodomain of the type II TGF-β receptor by scanning-deletion mutagenesis and knowledge-based modeling. *FEBS Letters* **456**:79-84.

Heldin, C.H., Miyazono, K., and ten Dijke, P. (1997). TGF-β signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**:465-71.

Hinck, A.P., Archer, S.J., Qian, S.W., Roberts, A.B., Sporn, M.B., Weatherbee, J.A., Tsang, M.L.S., Lucas, R., Zhang, B.L., Wenker, J. and Torchia, D.A. (1996). Transforming growth factorβ1: three-dimensional structure in solution and comparison with the X-ray structure of transforming growth factor β2. *Biochemistry* **35**: 8517-8534.

Hogan, B.L.M. (1996). Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev*. **10**:1580-94.

Innis, C.A., Shi, J., and Blundell, T.L. (2000). Evolutionary trace analysis of TGF-β and related growth factors: implications for site-directed mutagenesis. *Protein Eng* **13**:839-47.

Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C., and Vakser, I.A. (1992). Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. USA* **89**:2195-2199.

Kingsley, D.M. (1994). The TGF-β superfamily: new members, new receptors, and new genetic tests of function in different organisms. *Genes Dev* **8**:133-46.

Kirsch, T., Sebald, W. and Dreyer, M.K. (2000). Crystal structure of the BMP-2-BRIA ectodomain complex. *Nat. Struct. Biol*. **7**:492-496.

Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. 24: 946-950.

Laskowaski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst*. **26**: 283-291.

Lin, H.Y., Wang, X.F., Ng-Eaton, E., Weinberg, R.A., and Lodish, H.F. (1992). Expression cloning of the TGF-β type II receptor, a functional transmembrane serine/threonine kinase. *Cell*. **68**:775-85.

Lichtarge, O., Bourne, H.R., and Cohen, F.E. An evolutionary trace method defines binding surfaces common to protein families. J.Mol.Biol. 1996; 257: 342-58.

Massagué, J. (1990). The transforming growth factor-β family. *Annu. Rev. Cell Biol*. **6**:597-641.

Massagué, J. (1998). TGF-β signal transduction. *Annu. Rev. Biochem*. **67**: 753-91.

Massagué, J., Blain, S.W., and Lo, R.S. (2000). TGFβ signaling in growth control, cancer, and heritable disorders. *Cell* **103**:295-309.

Massagué, J. (2000) How cells read TGF-β signals. *Nat Rev Mol Cell Biol* **1**:169-78.

Mathews, L.S., and Vale, W.W. (1991). Expression cloning of an activin receptor, a predicted transmembrane serine kinase. *Cell*. **65**:973-82.

Mehler, M.F., Mabie, P.C., Zhang, D., and Kessler, J.A. (1997). Bone morphogenetic proteins in the nervous system. *Trends Neurosci*., **20**:309-17.

Mittal, P.R.E., Pristle, J.P., Cox, D.A., McMaster, G., Cerletti, N. and Grutter, M.G. (1996). The crystal structure of TGF-β3 and comparison to TGF-β2: implications for receptor binding. *Prot. Sci*. **5**: 1261-71.

Miyazono, K. (1993). Activin receptor-like kinases: a novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. *Oncogene*. **8**:2879-87.

Miyazono, K., ten Dijke, P., and Heldin, C.H. (2000). TGF-β signaling by Smad proteins. *Adv Immunol* **75**: 115-57.

Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. (1998). JOY: protein sequence-structure representation and analysis. *Bioinformatics* **14**: 617-623.

Murray-Rust, J., McDonald, N.Q., Blundell, T.L., Hosang, M., Oefner, C., Winkler, F., and Bradshaw, R.A. (1993). Topological similarities in TGF-β2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* **1**:153-9.

Nicholls, A., Sharp, K.A. and Honig, B. (1993). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Biophys. J*. **64**: 166-170.

Overington, J.P., Zhu, Z.Y., Sali, A., Johnson, M.S., Sowdhamini, R., Louie, G.V. and Blundell, T.L. (1993). Molecular recognition in protein families: A database of aligned three-dimensional structures of related proteins. *Biochem. Soc. Trans*. 21: 597-604.

Padgett, R.W., Das, P., and Krishna, S. (1998). TGF-β signaling, Smads, and tumor suppressors. *Bioessays*, **20**:382-90.

Qian, S.W., Burmester, J.K., Tsang, M.L.S., Weatherbee, J.A., Hinck, A.P., Ohlsen, D.J., Sporn, M.B., and Roberts, A.B. (1996). Binding affinity of transforming growth factor-β for its type II receptor is determined by the C-terminal region of the molecule.

J.Biol.Chem. **271**: 30656-662.

Raftery, L.A., and Sutherland, D.J. (1999). TGF-β family signal transduction in Drosophila development: from Mad to Smads. *Dev Biol*, **210**:251-68.
Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J.Mol.Biol*., **7**: 95-99

Ramachandran, G.N. and Sasisekharan,V. (1968). Conformation of polypeptides and proteins. *Adv. Protein. Chem*. **23**: 283-438.

Roberts, A.B., and Sporn, M.B.(1993). Physiological actions and clinical applications of transforming growth factor-β (TGF-β). *Growth Factors* **8**: 1-9.

Ruppert, R., Hoffmann, E. and Sebald,W. (1996). Human bone morphogenetic protein 2 contains a heparin-binding site which modifies its biological activity. *Eur. J. Biochem*. **237**: 295-302.

Saarma, M. (2000). GDNF - a stranger in the TGF-β superfamily? *Eur J Biochem* **267**: 6968-71.

Sali, A. and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J.Mol.Biol*. **234**: 799-815.

Sali, A. and Blundell, T.L. (1990). Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J. Mol. Biol*. **212**: 403-428.

Sayle, A.R. and Milner-White, E.J. (1995). RASMOL: biomolecular graphics for all. *T.I.B.S*. **20**: 374-376.

Scheufler, C., Sebald, W. and Hulsmeyer, M. (1999). Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution. *J. Mol. Biol*. **287**: 103-115.

Sowdhamini, R., Burke, D.F., Huang, J.F., Mizuguchi, K., Nagarajaram, H.A., Srinivasan, N., Steward, R.E., and Blundell, T.L. (1998). CAMPASS: a database of structurally aligned protein superfamilies. *Structure*. **6** :1087-94.

ten Dijke, P., Ichijo, H., Franzen, P., Schulz, P., Saras, J., Toyoshima, H., Heldin, C.H., Schlunegger, M. P. and Grutter, M.G. J. (1993). Refined crystal structure of human transforming growth factor β2 at 1.95A resolution. *J Mol. Biol*. **231**: 445-458.

Thomson, J.D., Gibson, T.J., Planwniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**: 4876-82.

Vakser, I.A. (1995). Protein docking for low-resolution structures. *Protein Eng*. **8**:371-377.

Vakser, I.A. (1996). Long-distance potentials: An approach to the multiple-minima problem in ligand-receptor interaction. *Protein Eng*. **9**:37-41.

Wuytens, G., Verschueren, K., de Winter, J.P., Gajendren, N., Beek, L., Devos, K., Bosman, F., de Waele, P., Andries, M., van den Eijnden-van Raaij, a.J.M., Smith, J.C., and Huylebroeck, D. (1999). Identification of two amino acids in activin A that are important for biological activity and binding to the activin type II receptors. *J.Biol.Chem.* **274**: 9821-9827.

Young, L., Jernigan, R.L. and Covell, D.G. (1994). A role for surface hydrophobicity in protein-protein recognition. *Prot. Sci*. **3**:717-729.

Zimmerman, C.M., and Padgett, R.W. (2000). Transforming growth factorβ signaling

mediators and modulators. *Gene* **249**:17-30.