

General and Specific Receptor Tyrosine Kinase Signaling Mechanisms Revealed by  
Analysis of Chimeric EGFR and IR Family Receptors

by  
Jacqueline M. McCabe

A dissertation submitted to Johns Hopkins University in conformity with the  
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland  
July 2015

## **Abstract**

Receptors embedded in the plasma membrane serve to relay signals from the outside to the inside of the cell and regulate cell growth and differentiation. Disruption of these signals can occur by mutation or abnormal expression of receptor proteins. These disruptions can result in aberrant cell growth and are found in many cancers. As a result, cell surface receptors are common targets of directed small molecule and biologic therapeutics. Two such targeted receptors are Smoothed (Smo) and the Epidermal Growth Factor Receptor (EGFR). Mutations in Smo are found in basal cell carcinoma, meningiomas and medulloblastoma. Overexpression or mutations of EGFR are found in glioblastoma and non-small cell lung cancer. Approved therapeutics targeting Smo and EGFR are currently used in the clinic, but acquired resistance to both Smo and EGFR targeted therapeutics generally occurs within months or at most a few years. More precisely defining the molecular mechanisms of activation and inhibition of these two receptors is of both fundamental interest as well as critical to understanding and improving current targeted therapeutics.

Ligand-induced dimerization of receptor tyrosine kinases (RTKs) is a generally accepted mechanism of activation for all classes of RTKs except the Insulin Receptor (IR) family. In contrast to all other RTK families, IR family members exist as preformed disulfide linked dimers on the cell surface. Recent evidence suggests that the ectodomain (ECD) of IR, and its homolog the type 1 insulin-like growth factor receptor (IGF1R), maintains separation of subunit transmembrane regions in the absence of ligand. Ligand binding releases this ectodomain autoinhibition and allows the transmembrane domains to come together and the intracellular kinase domains to transphosphorylate. Parallels

between the IR/IGF1R mechanism of activation and ligand-induced dimerization suggested that the IR/IGF1R ECDs may be able to regulate other families of RTKs, and we set out to understand how the IR/IGF1R activation mechanism relates to EGFR activation. We find that the IR/IGF1R ECD fails to activate the kinase domain of EGFR in either IR-EGFR or IGF1R-EGFR chimeric receptors. In contrast, the EGFR ECD is able to activate the IR kinase domain in an EGFR-IR fusion in an EGF-dependent fashion. The IR and EGFR mechanisms of activation thus appear to share some but not all features.

Despite years of intensive study and its central importance to animal development and human health, our understanding of the Hedgehog (Hh) signaling pathway remains riddled with gaps, question marks, assumptions, and poorly understood connections. In particular, understanding how interactions between Hh and Patched (Ptc), a 12-pass integral membrane protein, lead to modulation of the function of Smo, a 7-pass integral membrane protein, has defied standard biochemical characterization. Recent structural and biochemical characterizations of Smoothened domains have begun to unlock this riddle, however, and point towards improved cancer therapies. We present purification and biochemical characterization of near full-length Smo protein in the presence of small molecules targeting multiple Smo domains. These studies were carried out with the aim of developing the tools necessary to understand the molecular mechanisms governing Smo activity.

Advisor: Daniel J. Leahy

Reader: L. Mario Amzel

## **Acknowledgements**

A sincere and heartfelt thank you to all who have made the words on these pages, and the journey to them, meaningful. I am both humbled and grateful to have spent the past five years uncovering ever small and ever fascinating details of biology.

To my mentor, Daniel J. Leahy: Thank you for your example of patient leadership and scientific excellence. I am humbled to have spent five years learning and growing as a scientist and as a person while working in your laboratory. I take with me an incredible respect and love for basic science research. To the BCMB program and Biophysics Department at Johns Hopkins: Your collegial nature is an example par excellence for all research programs. To Paul D. Roepe at Georgetown University: Thank you for opening your lab to a young undergraduate, where I first learned, by example, how a parasite could be endlessly fascinating and what it takes to be a careful and rigorous biochemist.

To my parents, Gary and Donna: Thank you for buying me every book I ever asked for and unquestionably supporting my academic aspirations at each stage of my education. To my siblings and first examples of the power of genetic diversity: Thank you for keeping me grounded, laughing at and with me throughout this journey. To my family: Thank you for the countless examples of loyalty, expressions of confidence, and for being the reason I'll never live too far from home.

To my friends: Thank you for not being scientists and still asking how my cells were doing. For every visit, phone call, vacation and staycation, you made the past five years incredibly fun and are the reason the years flew by incredibly fast.



And most importantly, to Pat: Thank you for the past seven years, and for dreams of the future. I feel as though two names should go on this diploma. Your disarming laugh may be the number one contributing factor to any of my success here and in life.

Cheers.

## Table of Contents

|   |      |
|---|------|
| Abstract.....   | ii   |
| Acknowledgements.....   | iv   |
| List of Figures.....  | viii |
| List of Tables.....   | x    |
| Chapter 1. How IGF1 Activates its Receptor.....   | 1    |
| Acknowledgements.....   | 1    |
| Abstract.....   | 1    |
| Introduction.....   | 2    |
| Materials and Methods.....  | 5    |
| Results.....  | 19   |
| Discussion.....   | 28   |
| Chapter 2: General and Specific Receptor Tyrosine Kinase Signaling Mechanisms.....      | 54   |
| Revealed by Analysis of Chimeric EGFR and IR Family Receptors                           |      |
| Abstract.....   | 54   |
| Introduction.....   | 54   |
| Materials and Methods.....  | 58   |
| Results.....  | 62   |
| Discussion.....   | 65   |
| Chapter 3: Smoothened Structure and Function.....                                       | 76   |
| Acknowledgements.....   | 76   |
| Abstract.....   | 76   |
| Introduction.....   | 77   |
| Methods.....  | 90   |
| Results.....  | 92   |
| Discussion.....   | 94   |
| Chapter 4: Detergent-solubilized Patched Purified from Sf9 Cells Fails to Interact..... | 104  |
| Strongly With Cognate Hedgehog or Ihog Homologs   |      |
| Acknowledgements.....   | 104  |
| Abstract.....   | 104  |
| Introduction.....   | 105  |
| Methods.....  | 108  |
| Results.....  | 117  |

|                               |     |
|-------------------------------|-----|
| Discussion.....               | 124 |
| Bibliography.....             | 137 |
| <i>Curriculum Vitae</i> ..... | 155 |

## List of Figures

|             |  |    |
|-------------|--|----|
| Figure 1.1  | The extracellular domain of IGF1R autoinhibits IGF1R activity.....                     | 34 |
| Figure 1.2  | Residues on IR important for ligand binding.....                                       | 36 |
| Figure 1.3  | Phosphorylation of the ICD.....  | 37 |
| Figure 1.4  | Conserved IR ECD dimer interaction.....  | 38 |
| Figure 1.5  | Ligand-induced conformational change in the IR ECD.....                                | 40 |
| Figure 1.6  | The IGF1R ECD maintains TM separation.....   | 42 |
| Figure 1.7  | The TMs associate in IGF1R ECD- $\Delta$ L1-TM-fp.....                                 | 44 |
| Figure 1.8  | IGF1R TMs associate.....   | 45 |
| Figure 1.9  | TMs associate in active IGF1R-fl.....  | 47 |
| Figure 1.10 | IGF1 stimulates autophosphorylation.....   | 48 |
| Figure 1.11 | Surface mutations on the IGF1R kinase domain disrupt receptor.....<br>activation.      | 49 |
| Figure 1.12 | Buried surface analysis of crystal lattice pairs of IR and IGF1R kinase...<br>domains. | 50 |
| Figure 1.13 | Surface analysis of IGF1R kinase domain.....   | 51 |
| Figure 1.14 | Analysis of the autophosphorylation of IGF1R kinase clusters.....                      | 52 |
| Figure 1.15 | Model of IGF1R activation.....   | 53 |
| Figure 2.1  | Epidermal growth factor receptor activation.....                                       | 71 |
| Figure 2.2  | Transiently transfected IGF1R-EGFR chimeric receptors fail to activate.....            | 74 |
| Figure 2.3  | IGF1R and EGFR activation in CHO cell stable cell lines.....                           | 75 |
| Figure 3.1  | Major transmembrane components of Hh signal reception and.....<br>transduction.        | 96 |
| Figure 3.2  | Smo 7TM Region.....  | 97 |
| Figure 3.3  | Structure of class F GPCR CRDs.....  | 98 |

|            |  |     |
|------------|--|-----|
| Figure 3.4 | Smoothened-interacting small molecules.....                                    | 99  |
| Figure 3.5 | Purification of CRD-containing Smo constructs.....                             | 100 |
| Figure 3.6 | Purification of mSmo- $\Delta$ CRD- $\Delta$ CT.....                           | 101 |
| Figure 3.7 | Purification of CRD-mSmo- $\Delta$ CT.....                                     | 102 |
| Figure 3.8 | Thermal denaturation of purified Smo protein.....                              | 103 |
| Figure 4.1 | Functional Ptc expresses on the surface of Sf9 cells.....                      | 129 |
| Figure 4.2 | SEC of Ptc proteins after exchange into different detergents.....              | 130 |
| Figure 4.3 | Purification and characterization of Ptc proteins.....                         | 131 |
| Figure 4.4 | Pulldown of potential binding partners by MmPtcT1.....                         | 132 |
| Figure 4.5 | Pulldown of potential binding partners by DmPtcT1.....                         | 133 |
| Figure 4.6 | Co-expression and purification of DmPtcT1 with potential binding partners..... | 134 |
| Figure 4.7 | Co-expression and purification of MmPtcT1 with Shh.....                        | 136 |

## **List of Tables**

|           |                             |    |
|-----------|-----------------------------|----|
| Table 2.1 | RTK Expression Vectors..... | 72 |
| Table 2.2 | Panel of RTK Clones.....    | 73 |

## **Chapter 1. How IGF1 Activates Its Receptor**

### **Acknowledgements**

Chapter 1 contains the published manuscript by Kavran *et al*, 2014 in the journal *eLife*. My primary contribution was analyzing the importance of specific IGF1R kinase domain surfaces for receptor activation. I designed and cloned clusters of IGF1R kinase domain surface mutants. I established a cell based assay used to determine the role of these clusters in receptor activation (Figure 1.11). This work identified two distinct surfaces on the IGF1R kinase domain that disrupt receptor activation but showed that IGF1R does not activate via an EGFR-like asymmetric kinase dimer. The cell based assay was later adapted for use in cysteine substitution experiments probing association of IGF1R transmembrane regions during receptor activation.

### **Abstract**

The type I insulin-like growth factor receptor (IGF1R) is involved in growth and survival of normal and neoplastic cells. A ligand-dependent conformational change is thought to regulate IGF1R activity, but the nature of this change is unclear. We point out an underappreciated dimer in the crystal structure of the related Insulin Receptor (IR) with Insulin bound that allows direct comparison with unliganded IR and suggests a mechanism by which ligand regulates IR/IGF1R activity. We test this mechanism in a series of biochemical and biophysical assays and find the IGF1R ectodomain maintains an autoinhibited state in which the TMs are held apart. Ligand binding releases this constraint, allowing TM association and unleashing an intrinsic propensity of the intracellular regions to autophosphorylate. Enzymatic studies of full-length and kinase-

containing fragments show phosphorylated IGF1R is fully active independent of ligand and the extracellular-TM regions. The key step triggered by ligand binding is thus autophosphorylation.

## **Introduction**

Ligand binding to IR and IGF1R extracellular regions (ECDs) stimulates receptor kinase activity, leading to phosphorylation of multiple substrates and initiation of specific signaling cascades (Siddle, 2012). IR family members are unique among RTKs in forming constitutive dimers (of  $\alpha\beta$  subunits). Dimerization per se thus cannot be the activating signal, and activation is thought to involve a ligand-dependent conformational change (Frattali, 1992; Lemmon and Schlessinger, 2010). Part of the function of the ECD appears to be maintaining an inactive state in the absence of ligand as tryptic removal of the IR ECD results in constitutive activity (Tamura, 1983; Shoelson, 1988).

Two ligand-binding sites are present in each  $\alpha\beta$  dimer. Each site is composed of two distinct partial sites known as 'Site 1', which is composed of residues on L1 from one subunit and residues on the  $\alpha\text{CT}'$  helix of the other subunit (a prime is used to indicate the opposite subunit), and 'Site 2', which is composed of residues on Fn1' and Fn2' (Williams, 1995; Mynarcik, 1996; Whittaker, 2001, 2008; Smith, 2010) (Figure 1.2). A classic feature of ligand binding to IR and IGF1R is negative cooperativity (De Meyts, 2004), which implies communication between the two sites such that ligand binding to one site generates an asymmetric state of the receptor in which the affinity of the second site for ligand is weakened. The nature of this asymmetric state is not known, but



hydrogen–deuterium exchange experiments with IGF1R identified regions that are likely foci of this asymmetry (Houde and Demarest, 2011).

Crystal structures of the unliganded ECD of IR (McKern, 2006; Smith, 2010) and a fragment of the IR ECD bound to Insulin (Menting, 2013) have greatly aided efforts to understand the molecular mechanisms governing IR/IGF1R activity. In the absence of ligand, the  $\alpha\beta$  subunits of the ECD form a symmetric, antiparallel dimer shaped like an inverted ‘V’ (Figure 1.4A). The apex of this ‘V’ is formed by reciprocal interactions between L2-Fn1 domain pairs from opposing subunits (Figure 1.4C), with the Fn2 and Fn3 domains (Fn2–3) forming ‘legs’ with C-termini separated by 117 Å. The fragment of IR crystallized with Insulin encompassed the first four domains of the ECD (L1-CR-L2-Fn1) with the  $\alpha$ CT peptide fused to the C-terminus of Fn1. The structure of this fragment revealed the interaction between Site 1 and Insulin and rationalized a wealth of biochemical data. The nature of the activating conformational change in IR remained unclear, however, owing to the absence of the Fn2-ID-Fn3 regions, which contain Site 2, and the presence of an apparently nonphysiological dimer mediated by  $\alpha$ CT regions (Menting, 2013). Current models of IR/IGF1R activation envision only minor structural rearrangements of the ECD when ligand binds and posit a ligand-dependent twist of the Fn2–3 domains with little change in TM separation as the activating signal (Ward and Lawrence, 2012; Ward, 2013).

Although structural and mechanistic details of ligand-dependent activation of the receptor remain unclear, autophosphorylation of the kinase domain is a key element of receptor activation. Early studies with purified IR showed that autophosphorylation enhances IR kinase activity and that addition of Insulin led to a threefold to fivefold

increase in  $V_{\max}$  with no change in  $K_m$  for peptide substrate (Kasuga, 1983; Pike, 1986). More recent studies with isolated IGF1R kinase domains showed little phosphorylation of peptide substrates in the absence of autophosphorylation but increased activity with each successive trans-autophosphorylation of the three tyrosines in the kinase activation loop (Favelyukis, 2001) (Figure 1.3). Crystal structures of IR and IGF1R kinase domains in both inactive and active, phosphorylated conformations revealed canonical kinase domain features and rationalized how autophosphorylation promotes the active conformation (Hubbard, 1994; Hubbard, 1997; Favelyukis, 2001; Huse and Kuriyan, 2002; Munshi, 2002). A specific asymmetric dimer of kinase regions is necessary for activation of the Epidermal Growth Factor Receptor (EGFR) (Zhang, 2006), a related RTK, but whether or how kinase domain interactions play a role in regulating IR/IGF1R activity is not known.

To address the gaps in our understanding of the molecular mechanisms underlying IR family activation, we inspected crystal structures of the IR ECD for clues to the nature of conformational changes that occur when ligand binds. We identified an underappreciated dimer of the IR ECD fragment with Insulin bound that preserves key inter-subunit interactions present in the unliganded IR ECD and illuminates conformational changes that occur when ligand binds. In particular, an inter-subunit interaction that stabilizes the separation of the Fn2–3 ‘legs’ of the ECD is disrupted by ligand binding, suggesting that TM separation may be important for maintaining the receptor in an inactive state. We tested this model in a series of biochemical and biophysical assays and show that the IGF1R ECD indeed autoinhibits activity by holding the TMs apart in the absence of ligand. Ligand binding releases this autoinhibition and

allows the TMs to come together in an autophosphorylation-competent state. Enzymatic activity assays of purified, full-length IGF1R and several kinase-domain containing IGF1R fragments further show that, once phosphorylated, IGF1R is fully active independent of ligand or allosteric stimulation. The key step regulated by ligand binding is thus autophosphorylation and not kinase activity per se, and the role of the IR/IGF1R ECD is to inhibit activity in the absence of ligand rather than promote activity in the presence of ligand.

## **Materials and Methods**

### *Expression and purification of the IGF1R kinase domain (IGF1R-kin)*

A recombinant baculovirus was engineered to encode residues 956–1256 of IGF1R with the substitution C1207S using the pFastBacHT B plasmid (Life Technologies, Carlsbad, California). Sf9 cells at a concentration of  $2 \times 10^6$  cells/ml were infected with this P3 baculovirus and harvested 3 days post-infection. The purification protocol was based on a previously published procedure (Favelyukis, 2001). Cells were lysed by using a French pressure cell press in buffer containing 50 mM Tris pH 8, 50 mM NaCl, 0.2% Triton X-100, Benzonase nuclease (Sigma, St. Louis, Missouri), and protease inhibitor tablets (Roche, Germany). IGF1R-kin was isolated by immobilized metal affinity chromatography (IMAC), and the histidine-tag was removed by overnight incubation with TEV protease. Calf-intestinal phosphatase (CIP) (New England Biolabs, Ipswich, Massachusetts) was also included in the incubation to dephosphorylate the protein. The next day, protein was loaded onto a size-exclusion column (Superdex 75 26/60) equilibrated in 10 mM Hepes pH 7.5, 150 mM NaCl, and 5 mM  $\beta$ -ME, and the

monomeric fractions pooled. Pooled fraction were then loaded onto a monoQ column (5/50 GL) equilibrated in 10 mM Hepes pH 7.5, 5 mM  $\beta$ -ME and eluted with a linear gradient of NaCl. Fractions corresponding to the unphosphorylated protein, as judged by Coomassie Brilliant Blue stained native gel, were pooled, concentrated, and flash frozen in 20% glycerol. Phosphorylated fractions were pooled, concentrated, and incubated with 10 mM ATP and 30 mM  $\text{MgCl}_2$ , desalted, and then repurified on a mono Q column. Fractions corresponding to the fully phosphorylated IGF1R-kin, as judged by Coomassie stained native gel, were pooled, concentrated, and flash frozen in 20% glycerol.

*Expression and purification of the IGF1R juxtamembrane kinase region (IGF1R-jmk)*

IGF1R-jmk, residues 930–1256 of IGF1R with C1207S substitution, was expressed in Sf9 cells as described for IGF1R-kin. The purification protocol was based on a previously published method (Craddock, 2007). Cells were lysed using a French pressure cell in buffer containing 50 mM Tris pH 8.0, 1 M KCl, 10% glycerol, 0.2% Triton X-100, Benzonase, and protease inhibitor tablets. Clarified lysate was applied to an IMAC column (Biorad, Hercules, California) and eluted with imidazole. IGF1R-jmk was then incubated overnight with TEV and CIP to remove phosphorylation and the N-terminal histidine tag and dialyzed into 10 mM Tris pH 8.0, 200 mM KCl, 10% glycerol, 0.2% Triton-X100, 2.5 mM  $\beta$ -ME, 1 mM PMSF at 4°C. The cleavage reaction was reapplied to an IMAC column and the flow-through collected. The flow through was diluted to a final concentration of 50 mM KCl and applied to a HiTrap Q (GE Lifescience, Pittsburgh, Pennsylvania) column equilibrated in 10 mM Hepes pH 7.5, 0.2% Triton X-100, and 5 mM  $\beta$ -ME and eluted with a linear gradient of KCl. Fractions

containing IGF1R-jmk were pooled, diluted to a final concentration of 50 mM KCl, and loaded onto a monoQ column equilibrated in the same buffer as for the HiTrap Q column. IGF1R-jmk was eluted with a gradient of KCl. Fractions containing unphosphorylated IGF1R-jmk were analyzed as described above, concentrated, and flash frozen in 20% glycerol. Phosphorylated IGF1R-jmk fractions were pooled, concentrated, and incubated with 10 mM ATP, 30 mM MgCl<sub>2</sub>, and 10% glycerol. The sample was then desalted and purified via monoQ to isolate the fully phosphorylated IGF1R-jmk, which was then concentrated and flash frozen as described previously.

*Expression and purification of the IGF1R intracellular region (IGF1R-icd)*

A recombinant baculovirus was engineered to encode the entire IGF1R intracellular region, residues 930–1337 with the substitution C1207S and a C-terminal SBP tag, using the pFastBacHT B plasmid (Life Technologies). Expression, lysis, IMAC, and cleavage were the same as for IGF1R-jmk. Histidine-tagged *Yersinia enterocolitica* phosphatase (YOP) was used for dephosphorylation. After cleavage, IGF1R-icd was loaded onto an IMAC column and the flow-through was collected. Untagged IGF1R-icd was then loaded onto a Streptactin HiTrap column (GE Lifescience) and eluted with 10 mM Hepes pH 7.5, 150 mM KCl, 5 mM  $\beta$ -ME, and 2.5 mM desthiobiotin. This purification steps ensures that IGF1R-icd had an intact C-terminal tail. IGF1R-icd was then purified on a G75 size-exclusion column equilibrated in the same buffer lacking desthiobiotin. Monomeric IGF1R-icd was then pooled, diluted, and purified as described above to isolate unphosphorylated IGF1R-icd. To generate fully phosphorylated IGF1R-icd, partially phosphorylated IGF1R-icd was incubated with 10 mM ATP, 30 mM MgCl<sub>2</sub>, 0.05% Triton X-100, and 10% glycerol for 4 hr at 4°C. This sample was then

desalted into 10 mM Tris pH 8, 150 mM NaCl, 0.05% Triton X-100, concentrated, supplemented with 20% glycerol and flash-frozen.

*Expression and purification of full-length IGF1R (IGF1R-fl)*

IGF1R-fl, residues 1–1337 with a C-terminal histidine tag, was stably expressed in HEK293 GnTi– (ATCC, Manassas, Virginia) cells by Fugene transfection followed by FACS sorting. IGF1R-fl expressing cells were grown in a shaking incubator supplemented with 8% CO<sub>2</sub> in Freestyle Medium (Invitrogen, Carlsbad, California) supplemented with 1% FBS and 2 mM Glutamine, pelleted, and lysed in a liquid nitrogen mill (Spex Sample Prep, Metuchen, New Jersey). Lysed cells were resuspended in 50 mM Tris pH 8.0, 100 mM NaCl, 5 mM EDTA, 1 mM PMSF, 10% glycerol, and 2%  $\beta$ -octylglucoside for 1 hr at 4°C and spun at 200,000×g at 4°C for 1 hr. Clarified lysate was incubated with cyanogen bromide-resin (GE Healthcare) coupled to the monoclonal antibody 24–55 (Soos, 1992) for 1 hr. The resin was separated by centrifugation, washed sequentially in buffer containing 1 M NaCl, 0.5 M Urea, or 10 mM MgCl<sub>2</sub>. Protein was eluted with 25 mM Tris pH 8.0, 200 mM NaCl, 1 M MgCl<sub>2</sub>, 10% glycerol, and 2%  $\beta$ -octylglucoside, buffer exchanged on a desalting column into 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.1 mM PMSF, 10% glycerol, and 2%  $\beta$ -octylglucoside, and loaded onto a monoQ column equilibrated in 10 mM Tris pH 8.0, 1 mM EDTA, 0.1 mM PMSF, 2.5% glycerol, and 0.03% Triton X-100. IGF1R-fl was eluted with 300 mM NaCl, and IGF1R-fl containing fractions pooled, concentrated in a Pierce 100 kDa molecular weight cut-off spin concentrator, supplemented with 20% glycerol, and flash frozen. To produce dephosphorylated IGF1R-fl, the above protocol was followed with the addition of YOP to the resuspension step. To produce phosphorylated protein, the

above protocol was followed with the addition of 30 mM MgCl<sub>2</sub> and 10 mM ATP to the cell resuspension step.

#### *Expression and purification of IGF1*

A plasmid encoding an N-terminal histidine tagged version of IGF1 was transformed into BL21(DE3) Rosetta2 cells (EMD-Millipore, Billerica, Massachusetts). Cells were grown at 37°C in Terrific Broth to an OD<sub>600</sub> of 0.4, induced with 1 mM IPTG, and harvested 2 hr after induction. Cells were lysed using a French pressure cell press in 20 mM Tris pH 8.0 and 6 M Urea. Clarified supernatants were loaded onto an IMAC column and eluted into 12.5 mM Tris pH 8.0, 3 M Urea, and 125 mM imidazole. Elution fractions containing IGF1 were pooled and dialyzed three times: first for 2 hr at room temperature against 20 mM Glycine pH 10.5, 15 mM DTT, and 2 M Urea, second for 18 hr against 20 mM Glycine pH 10.5, 3 mM DTT, 1 M Urea, 1 M NaCl, 0.5 μM CuCl<sub>2</sub>, and 20% Ethanol, and third for 2 hr at 4°C against 25 mM Tris pH 8, 200 mM NaCl. The sample was centrifuged to remove insoluble aggregates, and the supernatant applied to a G75 sizing column equilibrated in 5 mM Tris pH 8, 200 mM NaCl. Fractions containing monomeric IGF1 were pooled, concentrated, and flash frozen in 50% glycerol.

#### *Purification of the 24–55 anti-IGF1R antibody*

A hybridoma cell line expressing monoclonal antibody 24–55 (Soos, 1992) (a generous gift from K Siddle) was grown in a shaking incubator in hybridoma medium (Gibco, Carlsbad, California). Conditioned medium was applied to a Protein G column, and the antibody eluted with 100 mM Glycine pH 2.7 and neutralized with Tris pH 9.0. The 24–55 antibody was concentrated and dialyzed into 100 mM NaHCO<sub>3</sub> and 500 mM NaCl and coupled to cyanogen bromide-resin according to manufacturer's directions.

### *Expression and purification of fluorescent proteins*

Plasmids encoding N-terminally hexa-histidine tagged EYFP, mCherry, or mTurquoise were transformed into *Escherichia coli* BL21(DE3) cells. Cultures were grown in Luria Broth at 37°C to OD<sub>600</sub> ~0.6, and the temperature dropped to 18°C for 30 min. IPTG was added to a final concentration of 1 mM, and the culture was grown overnight (~16 hr). Cells were harvested by centrifugation and lysed using a French pressure cell press. An IMAC column was loaded with clarified lysate, washed with 20 mM imidazole, and eluted with 250 mM imidazole. Protein-containing fractions were pooled, dialyzed, concentrated, and flash frozen. To minimize photobleaching, all expression and purification steps were performed in the dark. The purity of the fluorescent proteins was >90% after the nickel column as judged by Coomassie Brilliant Blue stained SDS-PAGE.

### *In vitro kinase assays*

Radiometric kinase assays to determine kinetic parameters were performed as described previously (Qiu, 2009). Reactions were performed in 25 µl of a buffer containing 50 mM HEPES pH 7.5, 50 mM NaCl, 10% Glycerol, 0.5 mM DTT, 0.2% Triton X-100, 10 mM MgCl<sub>2</sub>, 0.2 mg/ml ovalbumin, 0.1 mM Na<sub>3</sub>VO<sub>4</sub>, and either fixed or varying concentrations of ATP and Biotin-KKEEEEYMMMMG as the peptide substrate. A previously identified consensus sequence was used as the peptide substrate with an extra N-terminal lysine added to help with solubility (Songyang, 1995). Reactions were performed at 30°C and quenched by the addition of 10 µl of 100 mM EDTA. 10 µl of 20 mg/ml avidin (ThermoFisher, Waltham, Massachusetts) was added to the reactions and the mixture was transferred to a concentrator with a 30-kDa molecular



weight cutoff (Pall, Port Washington, New Jersey). Samples were then washed three times in buffer containing 0.5 M NaCl and 0.5 M Na<sub>2</sub>HPO<sub>4</sub>, pH 8.3 and counted. Turnover of the limiting substrate was less than 10%. Reactions were performed in duplicate, and values were generally within 20%. Data were fit by non-linear least-squares to using the Michaelis–Menten equation to obtain apparent  $K_m$  and  $k_{cat}$  values.

#### *In vitro autophosphorylation assays*

10 nM of purified IGF1R-fl was incubated with 2 mM ATP in the same buffer as the in vitro kinase assays either without or with 50 nM IGF1. Upon addition of ATP, the reactions were incubated at 30°C, and samples were removed at various time points and immediately quenched with 83  $\mu$ M EDTA. Samples were separated by denaturing SDS-PAGE electrophoresis and transferred to PVDF membranes. Blots were probed with an anti-phosphotyrosine (pY) (4G10, EMD-Millipore) or anti-IGF1R $\beta$  antibodies, scanned with a Typhoon Scanner (GE Lifesciences), and intensities quantified with ImageJ. For autophosphorylation of IGF1R-kin variants, the same assay was used except IGF1R proteins were at 1  $\mu$ M and ATP at 1 mM.

#### *Cell-based IGF1R activity assays*

IGF1R cell-based assays were based on established procedures (Liu, 2012). HEK293 cells (ATCC) were maintained in adherent culture in DMEM:F12 supplemented with 5% FBS. HEK293 cells were chosen due to low background endogenous IGF1R as judged by Western blot and equal transfection efficiency of IGF1R wild-type and kinase-inactive (D1205A/N) when compared with transient transfection of *igf1r* (–/–) mouse embryonic fibroblasts (a generous gift from R Baserga). HEK293 cells were plated in a six-well plate at  $1 \times 10^6$  cells/well and transiently transfected with IGF1R expression

plasmids using polyethylenimine (PEI, linear MW 25,000; Polysciences, Inc., Warrington, Pennsylvania) at an optimized ratio of 1  $\mu$ g DNA: 3  $\mu$ g PEI. Variant IGF1R genes were subcloned into pSSX-F, a version of pSGHV0 modified to eliminate the growth hormone tag and add a C-terminal Flag tag (Leahy, 2000). After 18 hr, cells were washed three times with 2 ml Ham's F12 supplemented with 1 mg/ml BSA and incubated in this medium for 3 hr at 37°C to serum starve. In designated wells, 20 nM IGF1 was added and incubated for 30 min at 37°C.

For cross-linking assays and assays using truncated IGF1R variants, the wells were washed with ice-cold phosphate buffered saline and then lysed for 30 min at 4°C in 250  $\mu$ l of RIPA buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) supplemented with 1 mM activated Na<sub>3</sub>VO<sub>4</sub>, 1 mM PMSF, Benzamide hydrochloride (Sigma), and 10 mM iodoacetamide to prevent further disulphide-bond formation during lysis (Cao, 1992). Lysates were clarified and samples were separated by either denaturing reducing or non-reducing SDS-PAGE electrophoresis. Western blot analysis was performed using anti-IGF1R $\beta$  (Santa Cruz Biotechnologies, Dallas, Texas) Western blots were developed using ECL2 (Thermo Scientific) and scanned using a Typhoon Imager. Bands corresponding to phosphorylated IGF1R from the 4G10 Western blot were quantified using ImageJ and normalized to receptor expression from quantified bands of the IGF1R- $\beta$  Western blot for each experiment.

For autophosphorylation analysis, the wells were washed with ice-cold phosphate buffered saline and lysed in supplemented RIPA buffer lacking iodoacetamide. The cell suspension was clarified, and total protein concentration determined using the BCA

protein assay (Thermo Scientific-Pierce) to normalize cell lysates. For assays with truncated IGF1R proteins, Western blot analysis was performed on cell lysates using anti-IGF1R $\beta$  or anti-IGF1R phosphotyrosine 1135 (pY1135) (Cell Signaling Technology, Danvers, Massachusetts) antibodies. For assays with IGF1R-fl proteins with kinase clusters, single cysteine substitutions, or TM mutations immunoprecipitation was performed. Anti-Flag-M2 (Sigma) was added at 0.5  $\mu$ g/ml to lysate followed by the addition of 20  $\mu$ l Protein G Sepharose 4 Fast Flow (GE Healthcare). Lysates were then incubated overnight at 4°C. Beads were then washed three times with 1 ml of RIPA buffer supplemented with 1 mM activated Na<sub>3</sub>VO<sub>4</sub>. Beads were eluted by the addition of 20  $\mu$ l of 5 $\times$  LDS loading buffer containing 10% fresh  $\beta$ -mercaptoethanol. Equal amounts of eluted proteins were analyzed as described for the cross-linking analysis but using both anti-pY and anti-IGF1R $\beta$  antibodies.

#### *Two photon FRET microscopy on living cells*

We employed two donor–acceptor FRET pairs for this study: mTurquoise-YFP and YFP-mCherry. We calculated the Förster radius for each pair to be 54.5 Å (mTurquoise-EYFP) and 53.1 Å (EYFP-mCherry), making them suitable for investigating conformational changes on the order of 30–90 Å. CHO cells were cultured in 35-mm collagen-coated glass bottomed dishes (MatTek Corporation, Ashland, Massachusetts) using phenol red-free DMEM-F12 supplemented with 5% FBS and 1 mM L-glutamine. Plasmids encoding the IGF1R extracellular and transmembrane regions fused after GGSGGS to mTurquoise (FRET donor) or EYFP (FRET acceptor) were co-transfected using Fugene HD (Promega, Madison, Wisconsin) at a mass ratio of 3:1

(Fugene:DNA), according to the manufacturer's protocol. Transfection proceeded for 24 hr at 37°C, 5% CO<sub>2</sub>.

We used the OptiMiS TruLine Spectral Scanning System (Aurora Spectral Technologies, Milwaukee, Wisconsin) for two-photon microscopy. A solid-state continuous wave laser pumped a mode-locked Ti:Sapphire laser (MaiTai DeepSee, Newport Corporation, Irvine, California) that generated near infra-red pulses in the wavelength range of 800–960 nm ( $\lambda_{exc}$  for mTurquoise = 800 nm;  $\lambda_{exc}$  for eYFP = 960 nm). The beam was focused using a 63× objective (Nikon, Japan). Fluorescence emission from the sample chamber was projected through a transmission grating onto a cooled electron-multiplying charge-coupled device. The full spectral scans of all pixels in the viewing field (300 × 440 pixels) were completed in about 10 s.

Image acquisition and processing are based on the method described by (Raicu, 2008). Prior to imaging, CHO cells were washed three times in 1× PBS to remove all traces of FBS, then serum starved in phenol red-free DMEM-F12 without FBS. After 5 hr of starvation, the media were changed a final time and the cells were imaged directly in the two-photon microscope. Individual donor and acceptor fluorescence spectra were collected from cells expressing either the donor- or acceptor-fused IGF1R alone. A single ‘FRET’ scan ( $\lambda_{exc}$ , two photon = 800 nm) was acquired for each cell that coexpressed both EYFP and mTurquoise-fusion proteins. The full spectral emission profile was collected from 450–600 nm (1 nm resolution), and the FRET efficiency per pixel ( $E_{APP}^D$ ) was calculated using a Matlab program, according to the equations described (Raicu, 2008)

$$E_{app}^{Dq} = \frac{F^D(RET)}{F^D(\lambda_{ex})} = \frac{1}{1 + \frac{Q^A}{Q^D} \frac{k^{DA}}{k^{AD}} \frac{w^D}{w^A}}$$

where  $F^D(RET)$  and  $F^D(\lambda_{ex})$  equal the fluorescence emission of the donor in the presence and absence of resonance energy transfer, respectively.  $Q^A$  and  $Q^D$  are the quantum yields of the donor and acceptor fluorophore, and  $k^{DA}$  and  $k^{AD}$  are the maximum emission intensities of the donor in the presence of the acceptor, and the acceptor in the presence of the donor, respectively. The integrals of the individual emission spectra for the donor and acceptor are given by  $w^D$  and  $w^A$ . Representative pixels in each cell membrane were examined by eye to ensure a good fit to the FRET model. We calculated the Förster radius for the mTurquoise-EYFP donor–acceptor pair to be 54.5 Å using measured absorption and emission spectra from purified proteins (King, 2014).

#### *Confocal microscopy and quantitative imaging FRET*

CHO cells were cultured in DMEM-F12 supplemented with 5% FBS and 1 mM L-glutamine. Cells were seeded in 35-mm dishes at a density of  $4 \times 10^4$  cells per well and grown for 24 hr at 37°C in 5% CO<sub>2</sub>, then transiently co-transfected with plasmids encoding IGF1R TM-EYFP (3 µg) and IGF1R TM-mCherry (6 µg) using the Fugene HD transfection reagent at a mass ratio of 3:1 (Fugene:DNA) according to the manufacturer's protocol.

After 24 hr, cells were washed twice with 1 ml of 30% PBS, one minute per wash. 1 ml of chloride salt vesiculation buffer (Del Piccolo, 2012) was added to each well, and the wells were incubated at 37°C, 5% CO<sub>2</sub>. Vesiculation buffer is composed of 200 mM NaCl, 5 mM KCl, 0.5 mM MgCl<sub>2</sub>, 0.75 mM CaCl<sub>2</sub>, and 100 mM bicine, pH 8.5. Vesiculation reached completion after about 12 hr, and the entire well supernatant was

transferred to 4-well chambered slides (Thermo Scientific, Nunc Lab-Tek II). The wells were allowed to equilibrate to room temperature before imaging.

Vesicle images were acquired using a Nikon C1 laser scanning confocal microscope equipped with a 60× water immersion objective. Three scans were taken for each vesicle: (i) a ‘donor’ scan ( $\lambda_{exc} = 488 \text{ nm}$ ,  $\lambda_{em} = 500\text{--}530 \text{ nm}$ ), (ii) an ‘acceptor’ scan ( $\lambda_{exc} = 543 \text{ nm}$ ,  $\lambda_{em} = 650 \text{ nm longpass}$ ), and (iii) a ‘FRET’ scan ( $\lambda_{exc} = 488 \text{ nm}$ ,  $\lambda_{em} = 565\text{--}615 \text{ nm}$ ). The donor and FRET scans used a 488 nm argon ion laser excitation source, while the acceptor scan used a 543 nm He-Ne laser. The image resolution was  $512 \times 512$  pixels, with a pixel dwell time of  $1.68 \mu\text{s}$ . The gains were set to 7. All images were processed using a Matlab program developed in the Hristova laboratory. The program finds the boundary of each vesicle, verifies that the vesicle is present in all three scans, and fits the intensity profile across the membrane to a Gaussian function. The baseline is fitted with an error function. Details for the calculations of FRET efficiency are provided in the study by Chen et al. (2010a, 2010b). We calculated the intrinsic FRET efficiency as:

$$\text{Intrinsic FRET} = (E_{app} - E_{proximity}) / x_a$$

where  $E_{app}$  is the apparent FRET efficiency observed in each vesicle,  $E_{proximity}$  is the contribution of FRET from nonspecific interactions (King, 2014), and  $x_a$  is a correction factor which accounts for varying ratios of donor and acceptor molecules within each vesicle.

To determine fluorescent protein concentration in the vesicles, vesicle fluorescence intensities were normalized to standard solutions of mCherry and EYFP purified from *E. coli* BL21 cells using the formula (Chen, 2010a, 2010b):

$$E + 1 - \frac{I_D^m}{I_D^m + I_{D,corr}^m}$$

where  $I_D^m$  is the emission intensity of the donor per unit area of the membrane in the presence of RET, and  $I_{D,corr}^m$  is the donor emission intensity in the absence of RET (Chen, 2010a, 2010b). Bleed through coefficients were calculated from confocal images of standard solutions of purified EYFP and mCherry. Using absorption and emission spectra of purified proteins, we calculated the Förster radius for the EYFP-mCherry donor–acceptor pair to be 53.1 Å.

### *Molecular dynamics simulations*

A simulation system was set up by tiling an equilibrated 15% POPS–85% POPC bilayer structure to a rectangular parallelepiped box of desired dimensions. The simulation box was filled with water molecules and 0.15 M NaCl. The POPS fraction was chosen to mimic the abundance of anionic lipids in mammalian plasma membrane (Zachowski, 1993; van Meer, 2008), in which ~10% of lipids are PS species, and few more percent are other anionic species, such as phosphoinositides. The bilayer was modeled with POPS only in the intracellular leaflet. The IGF1R TM helices were embedded in the bilayer. The simulation box of two IGF1R TM helices for the two simulations of 12  $\mu$ s and 34  $\mu$ s is of 65 Å  $\times$  65 Å  $\times$  86 Å in dimensions and the simulation box of four IGF-1R transmembrane helices for the other two simulations of 24  $\mu$ s and 36  $\mu$ s is of 81 Å  $\times$  81 Å  $\times$  93 Å in dimensions.

The systems were parameterized using the CHARMM36 forcefield for lipids and salt (Beglov and Roux, 1994), CHARMM TIP3P for water molecules (Neria, 1996), and the CHARMM22\* forcefield (Mackerell, 1998, 2004; Piana, 2011) for protein. After

parameterization, the systems were energy minimized and were equilibrated in a 10 ns simulation with a NPT ensemble at 310 K temperature and 1 bar pressure. In the equilibration, 1 fs time step was used and harmonic positional restraints of a force constant of 1 kcal/mol/Å<sup>2</sup> were initially applied to the protein backbone atoms.

A production simulation was launched using the last frame of the equilibration simulation, with a NPT ensemble at 310 K and 1 bar using the Berendsen coupling scheme (Berendsen, 1984) with one temperature group and Nosé-Hoover thermostat (Hoover, 1985) with a time constant of 0.05 ps. All hydrogen-containing bonds were constrained using a recent implementation (Lippert, 2007) of the M-SHAKE algorithm (Krautler, 2001). Reversible reference system propagator algorithm (r-RESPA) integrator (Tuckerman and Berne, 1992) was used. Bonded, Van der Waals, and short-range electrostatic interactions were computed every 2 fs, while long-range electrostatic interactions were computed every 6 fs. All molecular dynamics simulations were performed using Anton, a special purpose supercomputer (Shaw, 2009) designed for such simulations. The short-range electrostatic interactions were calculated at a cutoff of 9.48 Å, and the long-range electrostatic interactions were computed using Gaussian-split Ewald (Shan, 2005).

### *Structural analysis*

Crystal lattice contacts for each crystal form analyzed were identified manually using the program PyMol and each kinase pair was individually defined. The Lsqkab program in CCP4 was used to superpose the kinase molecules of each pair on each kinase molecule in every other pair to identify any conserved interactions. Buried surface areas for each pair were calculated using Areaimol in CCP4.



## Results

### *Ligand binding disrupts interactions stabilizing IGF1R/IR TM separation*

To search for ligand-dependent conformational changes in the IR ECD, we compared the crystal structures of liganded and unliganded IR ECD fragments (Smith, 2010; Menting, 2013). The reported dimer of the liganded IR fragment is mediated only by reciprocal interactions between L1 from one subunit and  $\alpha$ CT from the other. As this dimer neither resembles the inverted 'V' conformation of the unliganded IR ECD nor is compatible with the inter-subunit disulfide bond between Fn1 domains, it is likely a nonphysiological artifact arising from the deletion of Fn2-ID-Fn3 domains and fusion of  $\alpha$ CT to Fn1 (Menting, 2013). We therefore examined all IR interactions present in the crystal and identified an alternative dimer (Figure 1.4B), which had previously been noted as a part of a dimer of dimers (Menting, 2013), that is compatible with the disulfide bond between Fn1 domains (Schaffer and Ljungqvist, 1992) and preserves the inter-subunit contacts between opposing L2-Fn1 domains (L2-Fn1:L2'-Fn1') present at the apex of the inverted 'V' in the unliganded IR structure (Figure 1.4). Superposition L2-Fn1:L2'-Fn1' domains from the unliganded IR structure (Smith, 2010) and this symmetry-generated dimer of liganded IR (Menting, 2013) results in a root mean square deviation of 3.4 Å for 478 C $\alpha$  atoms (Figure 1.4C and 1.4D).

Using this superposition as a basis for comparing liganded and unliganded IR conformations reveals that the L1-CR domains rotate about a hinge between CR and L2 (Figure 1.5A and 1.5B) when ligand binds. The position of this hinge is homologous to that of the hinge observed in the ECD of EGFR when ligand binds, and most of the  $\sim 50^\circ$

rotation can be ascribed to a rotation about the psi angle of Gly 306 (Burgess, 2003; Menting, 2013). The rigid-body rotation of the L1-CR domains repositions them further from the central axis of the IR dimer, exposes several residues on L1 that mediate Insulin binding, and disrupts an extensive interface between L1 and Fn2'-3' domains in unliganded IR (Figure 1.5C and 1.5D). (Hubbard, 2013; Menting, 2013). This interface fixes the relative positions of the Fn2-3 tandems in unliganded IR and stabilizes the ~120 Å separation of the Fn3 domains (and the immediately adjacent TM regions) observed in the absence of ligand.

This observation led us to ask whether the TM separation enforced by the L1:Fn2'-3' interface maintains the receptor in an inactive state in the absence of ligand. To test this possibility, we substituted several L1 residues that are both conserved between IGF1R and IR and occur at the L1:Fn2'-3' interface with residues of opposite charge or increased size. These substitutions consistently resulted in poor processing and/or expression of IGF1R, however, which precluded analysis of their effects on IGF1R activation. We then analyzed an IGF1R variant lacking the entire L1 domain (IGF1R  $\Delta$ L1) and found that it expressed well in HEK293 cells and was constitutively phosphorylated as judged by anti-phosphotyrosine Western blots (Figure 1.1C) demonstrating that the L1 domain is essential for maintaining the inactive state of IGF1R.

#### *The IGF1R extracellular region is autoinhibitory*

As deletion of the L1 domain results in constitutive activity of IGF1R, we next tested whether the role of the ECD might be to stabilize an inactive state in the absence of ligand as opposed to stabilizing the active state in the presence of ligand. We deleted most or all of the IGF1R ECD and found that these truncated IGF1Rs (IGF1R Fn3-TM-

icd and IGF1R TM-icd) are constitutively phosphorylated in the absence of ligand (Figure 1.1D). A kinase-inactivating point mutation eliminated this phosphorylation, implicating autophosphorylation as the source of this signal. The IGF1R ECD thus plays an active role in keeping the receptor off in the absence of ligand but is not needed to activate the truncated receptor.

*Release of IGF1R autoinhibition by ligand brings the TMs together*

Loss of either the L1 domain or the ECD eliminates the inter-subunit interaction between L1:Fn2'–3' that fixes the separation of the TM regions. To examine if TM separation is altered in active receptors, we replaced the intracellular domain (ICD) of IGF1R with fluorescent donor or acceptor proteins (IGF1R ECD-TM-fp) and determined the FRET efficiency between co-transfected donor–acceptor pairs in the presence and absence of ligand (Figure 1.6A). We used spectrally resolved Förster resonance energy transfer (FRET) (Raicu, 2008) to calculate the FRET efficiency at the pixel level in transfected CHO cells, which ensures only proteins expressed at the cell membrane are included in the analysis. FRET efficiency increased twofold in the presence of ligand (Figure 1.6A). Assuming comparable if not random orientations of the fluorescent proteins in both states, these data indicate the C-termini of the TMs move closer together in the presence of ligand. A similar increase in FRET efficiency was observed in vesicles derived from CHO cells expressing IGF1R ECD-TM-fp in the presence and absence of ligand. We also measured the FRET efficiency for IGF1R lacking the L1 domain (IGF1R ECD $\Delta$ L1-TM-fp) and found that its FRET efficiency is similar to that of IGF1R ECD-TM-fp with bound ligand suggesting that the ligand bound state of IGF1R is similar to that of the receptor when the L1:Fn2'–3' interface is disrupted (Figure 1.7) by

enforcing TM separation, introduction of flexible linkers between the ECD and TM regions should release this inhibition, and insertion of progressively longer linkers composed of glycine–serine containing repeats between 4 and 20 residues long indeed led to increasing levels of basal receptor activity (Figure 1.6B).

#### *Isolated IGF1R TMs dimerize*

The ligand-dependent decrease in distance between IGF1R TMs observed in FRET assays of IGF1R ECD-TM-fp led to the question of whether IGF1R TMs physically associate in the active state. Analysis of IR family TM sequences failed to identify any motifs known to mediate TM interactions, such as GXXXG (Russ and Engelman, 2000), but did identify an absolutely conserved proline (P911 in human IGF1R) (Figure 1.8A). To investigate the role of P911 in receptor activation, we performed cell based activity assays with IGF1R-fl bearing either a single P911L substitution or with additional substitutions to residues surrounding P911. These IGF1R variants became phosphorylated in a ligand-dependent manner suggesting P911 does not play a key role in mediating receptor activity, an outcome consistent with the results of earlier studies involving more dramatic manipulations of IR TM residues and suggesting lax constraints on TM sequence (Frattali, 1991; Yamada, 1992; Cheatham, 1993).

To investigate the behavior of IGF1R TMs further, we performed molecular dynamics (MD) simulations of IGF1R TM monomers inserted into a 15% phosphatidylserine, 85% phosphatidylcholine bilayer. Four simulations extending between 12  $\mu$ s and 36  $\mu$ s were performed. In two of these simulations IGF1R TMs formed dimers in which the two TM helices cross each other opposite the conserved proline (Figure 1.8B). These dimers formed early in the simulations and were stable

throughout the rest of the runs. The IGF1R TMs also adopted several stable monomeric conformations in these simulations, occasionally with a sharp kink at P911 (Figure 1.8B).

Experimental evidence that IGF1R TMs have an intrinsic propensity to dimerize in bilayers was obtained from quantitative imaging FRET (QI-FRET) experiments (Li, 2008). Isolated IGF1R TMs were fused to C-terminal donor or acceptor fluorescent proteins (IGF1R TM-fp) and transiently transfected into CHO cells. Vesicles containing variable amounts of expressed proteins were generated and FRET efficiency analyzed (Chen, 2010a; Del Piccolo, 2012). FRET efficiency increased as a function of fluorescent protein concentration and exceeded both the theoretical levels of FRET expected to arise from random proximity (King, 2014) and the FRET efficiency of a control monomeric membrane protein (ErbB2 ECD-TM-fp) demonstrating an intrinsic propensity of IGF1R TMs to associate (Figure 1.8C). When examined on the surface of cells the FRET efficiency of IGF1R TM-fp overlapped with the FRET efficiency observed for IGF1R ECD-TM-fp in the presence of ligand suggesting similar TM separations in liganded IGF1R and isolated IGF1R TM dimers

An intrinsic propensity of IGF1R TM and ICD regions to dimerize was also demonstrated by cysteine substitutions in constitutively active ECD-truncated forms of IGF1R (IGF1R TM-icd) (Figure 1.8D). The side chains of the extracellular membrane-proximal residue H905 approach within 6 Å of one another in the MD dimer (Figure 1.8B), and H905 or T898 were individually substituted with cysteine in IGF1R TM-icd variants. Western blots of reduced and non-reduced lysates from HEK293 cells expressing these proteins show that the majority of the H905C variant forms a disulfide-linked dimer (Figure 1.8D). Lower amounts of dimerization were observed for the T898

variant in which the cysteines are farther away from the TM region (Figure 1.8D), and no disulfide crosslinking was observed for an equivalently truncated IGF1R without a cysteine substitution. These results demonstrate that, in the absence of the ECD, the IGF1R TM-icds come together in a manner consistent with the MD TM dimer.

*IGF1R TMs dimerize in the active, full-length receptor*

We next investigated whether juxtamembrane cysteines formed disulfide crosslinks when introduced into IGF1R-fl (Figure 1.9A). We could not directly analyze intra-dimer disulfide bond formation by introduced cysteines by analyzing migration on reducing and non-reducing SDS-PAGE owing to the native disulfide bonds linking  $\alpha$  and  $\beta$  chains. To circumvent this issue, we substituted the cysteine residues that mediate the  $\alpha\beta$  disulfide with serine. These substitutions inhibited proper proteolytic processing of the receptor, however, obscuring differentiation of the new disulfide bond from native disulfide bonds. To our surprise, however, Western blots monitoring the phosphorylation of the IGF1R variants with the single substitutions of either T898C or H905C revealed high levels of phosphorylation in the absence of ligand (Figure 1.9A). An IGF1R-fl variant with a cysteine substitution immediately following the Fn3 domain and further away from the TM (Q895C) displayed only a modest increase in basal phosphorylation compared to wild-type (Figure 1.9A), perhaps owing to steric interference from Fn3 domains.

The most likely explanation for the ligand-independent activity of IGF1R variants with single cysteine substitutions is that IGF1R samples the active conformation in the absence of ligand and becomes trapped in the active state by a disulfide bond. An intra-dimer disulfide bond could not occur at these positions if the receptor was locked in the inactive 'legs' apart state by the L1:Fn2'-3' interaction. This interaction must open

transiently for ligand to bind, however, and it appears this transient opening allows TM association. Alternatively, inter-dimer disulfides could form and activate the receptor. To distinguish these possibilities, we compared the migration of wild-type and IGF1R H905C on reducing and non-reducing SDS-PAGE. If the introduced cysteines form inter-dimer disulfide bonds, some fraction of IGF1R H905C would migrate as a dimer of dimers or higher (molecular weight  $\geq \sim 640$  kDa vs.  $\sim 320$  kDa for a dimer) on non-reducing SDS-PAGE, and the ratio of H905C dimer to total monomer would decrease relative to wild-type IGF1R. Western blots show that this ratio remains the same for H905C and wild-type IGF1R, however, indicating the absence of appreciable inter-dimer disulfide formation by IGF1R H905C. Sequence alignments show that H905, and the IGF1R juxtamembrane region in general, can be substituted with at least eleven different amino acids indicating the effects of cysteine at this site likely stem from its ability to form disulfide bonds.

An intra-dimer disulfide between cysteines at position 905 would force TM association. When the H905C substitution was introduced into IGF1R ECD-TM-fp, the FRET efficiency of this variant in the absence of ligand matched that of native IGF1R ECD-TM-fp in the presence of ligand (Figure 1.9B). This result shows that the H905C substitution indeed leads to decreased TM separation, presumably owing to crosslinked juxtamembrane regions, and that this conformation is indistinguishable from that observed for liganded IGF1R ECD-TM-fp.

#### *Ligand binding does not stimulate the kinase activity of phosphorylated IGF1R*

To understand how changes in IGF1R ECD conformation are linked to receptor kinase activity, we next assessed the effects of ligand binding and autophosphorylation

on the kinase activity of IGF1R. We expressed full-length IGF1R (IGF1R-fl) in HEK293 cells and purified it to near homogeneity for in vitro kinase assays. To parse the contributions of specific regions to regulation of IGF1R activity, we expressed the isolated kinase domain (IGF1R-kin), the N-terminal juxtamembrane segment linked to the kinase domain (IGF1R-jmk), and the entire intracellular domain (IGF1R-icd) in Sf9 cells and purified each fragment. Either fully phosphorylated or unphosphorylated (or minimally phosphorylated) versions of each form of IGF1R were prepared by incubation with ATP and  $Mg^{++}$  or phosphatase, respectively, during purification.

The steady-state kinetic parameters of each form of IGF1R were determined using a direct, radiometric assay monitoring phosphotransfer from [ $\gamma$ - $^{32}P$ ]ATP to a biotinylated peptide substrate. Activity in this assay was linear with respect to both time and enzyme concentration for all forms of IGF1R. Each intracellular IGF1R fragment showed barely detectable catalytic activity when unphosphorylated, which differed from the low but measurable activity of phosphatase-treated IGF1R-fl. Western blots indicated incomplete dephosphorylation of IGF1R-fl despite efforts to drive this reaction to completion, however, and residual autophosphorylation likely accounts for the elevated activity of IGF1R-fl relative to unphosphorylated IGF1R fragments.

In contrast, robust activity was observed for all phosphorylated forms of IGF1R.  $K_{m,app}$  values for both ATP and peptide ranged between 50 and 110  $\mu M$ , and  $k_{cat}$  values ranged between 98 and 749  $min^{-1}$ . Addition of the juxtamembrane region (IGF1R-jmk) resulted in fourfold increase in  $k_{cat}$  from that observed for IGF1R-kin, but addition of the C-tail (IGF1R-icd) had no effect on activity (Figure 1.10A). Notably,  $k_{cat}$  differences of less than twofold were observed between phosphorylated forms of IGF1R-fl and either



IGF1R-jmk or IGF1R-icd indicating that the activities of these phosphoproteins are essentially equivalent. That is, the C- tail, ECD, and TMs do not contribute to the activity of phosphorylated IGF1R. Surprisingly, the  $k_{cat}$  of phosphorylated IGF1R-fl is marginally less with ligand than without indicating that bound ligand also does not contribute to the activity of phosphorylated IGF1R-fl.

*Ligand binding stimulates IGF1R autophosphorylation*

Since the activity of phosphorylated IGF1R-fl is high and unaffected by ligand binding, we reasoned that autophosphorylation must be the step controlled by ligand-dependent allostery. To test this idea, phosphatase-treated IGF1R-fl was incubated with ATP and  $Mg^{++}$  in the presence and absence of IGF1, and the rate of autophosphorylation monitored by Western blotting. The presence of IGF1 indeed led to a marked increase in the rate of autophosphorylation (Figure 1.10B).

*No evidence for specific intracellular domain interactions*

Interactions between kinase domains have proven essential for stabilizing the active state of EGFR (Zhang, 2006), and we wondered if specific interactions between kinase domains are coupled to active or inactive states in the IR family. To address this question, we analyzed lattice contacts in crystals of active and inactive IGF1R and IR kinases to determine if any contacts appear repeatedly or bear hallmarks of a physiological interaction. Several of nearly 300 unique contacts identified bury more than 800 Å<sup>2</sup> of surface area (Bahadur, 2004) but none appeared in multiple crystal forms or suggest a rationale for maintaining a stable active or inactive state (Figure 1.12).

We next used mutagenesis and a cell-based activity assay to search for surfaces on the IGF1R kinase important for maintaining an inactive or active state. Six clusters of

3–9 alanine substitutions that blanket the IGF1R kinase surface were created (Figure 1.13). Sites selected for substitution are solvent exposed, conserved between IGF1R and IR, and not part of the active site. Tagged versions of all six IGF1R variants were expressed on the surface of HEK293 cells (Figure 1.11A) and their levels of phosphorylation were assessed by Western blot after immunoprecipitation. In the absence of ligand, all variant IGF1Rs showed little or no phosphorylation, similar to wild-type IGF1R, and provide no evidence for an autoinhibitory interaction between kinases (Figure 1.11A).

In contrast, two clusters of substitutions, one on the kinase N-lobe (C1) and one on the C-lobe (C4), resulted in loss of IGF1R phosphorylation in the presence of ligand (Figure 1.11A). Co-transfection of the two inactivating IGF1R N- and C-lobe variants did not restore activity indicating that they are unlikely to represent complementary surfaces of an asymmetric kinase dimer (Figure 1.11B). To ensure that the inactivating N- and C-lobe mutations did not affect the intrinsic kinase activity of the receptor, IGF1R-kin proteins bearing these mutations were expressed and purified. Both variants retained autophosphorylation and substrate kinase activities but at lower levels than wild-type IGF1R-kin (Figure 1.14) and it cannot be ruled out that this decrease in intrinsic activity underlies the loss of ligand-dependent autophosphorylation observed in cells.

## **Discussion**

We describe a molecular mechanism for IR/IGF1R activation in which ECD-enforced separation of TMs maintains the receptor in an inhibited state. Ligand binding relieves this inhibition by disrupting the L1:Fn2'–3' interaction that stabilizes TM

separation, freeing the TMs to associate and autophosphorylation of the kinase domains to proceed (Figure 1.15A). This model is based on a previously underappreciated dimer in the crystal structure of an Insulin bound IR ECD fragment, which provides a basis for comparing liganded and unliganded structures of the IR ECD.

To validate this model, we present FRET and mutagenesis studies that demonstrate (i) a decrease in IGF1R TM separation when ligand binds, (ii) an intrinsic propensity of the IGF1R TM and TM-icd to dimerize, (iii) that the kinase is active in this associated state, consistent with the increased activity of the IGF1R ICD when fused to a dimeric partner (Baer, 2001), (iv) that the role of the ECD is primarily to inhibit this intrinsic kinase activity in the absence of ligand, rather than promote a specific active state, (v) that IGF1R TMs associate in an active form of the full-length receptor, and (vi) that IGF1R samples the active state even in the absence of ligand. Enzymatic studies of full-length IGF1R and IGF1R fragments further demonstrate no role for allostery in maintaining IGF1R activity once it is phosphorylated, indicating that the key step regulated by ligand binding is autophosphorylation. Although ligand binding has no effect on the intrinsic activity of phosphorylated IGF1R, the short half-life of IGF1R phosphorylation when phosphatases are present suggests the continued presence of ligand may be needed to maintain IGF1R activity (Kleiman, 2011). This efficiency of cellular phosphatases may also explain why transient sampling of the active state by IGF1R does not lead to constitutive activity.

Previous models for IR activation have posited only modest changes in receptor conformation when ligand binds and that ‘twist’ rather than lateral movement of the Fn2–3 domains is the activating signal or that the TM regions move from associated to apart

during IR activation (Ward, 2013; Lee, 2014). We show here that ligand binding to IGF1R leads to TM association. Thus the Fn2–3 domains must undergo a large movement upon ligand binding to accommodate the new position of the TMs. To model this movement, we appended the Fn2–3 tandems from the unliganded IR ECD structure to the structure of the IR ECD fragment bound to Insulin. The Fn2–3 tandems were then manually repositioned using only the constraints that the ligand bound position of these ‘legs’ must be compatible with the TM dimer observed in MD simulations and that the residues on Fn2 that compose Site 2 contact ligand. Assuming the connection between Fn1 and Fn2 to be flexible (it buries  $\sim 500 \text{ \AA}^2$  of surface area) and the Fn2–3 tandem to be rigid, we find it possible to place the Fn2–3 tandems so that Site 2 residues appose ligand and that the ECD C-termini are compatible with TM dimerization (Figure 1.15B–D).

We have modeled IR/IGF1R with two bound ligands (Figure 1.15D) but receptors with a single ligand bound may represent an important active state. Negative cooperativity implies that ligand binding at one site influences assembly of the second site. The most likely conduit of information between sites is through the  $\alpha$ CTs, which are connected by an inter-subunit disulfide bond 8 residues from their N-termini. The  $\alpha$ CTs are 30  $\text{\AA}$  farther apart in the liganded IR structure than in the unliganded IR structure (Smith, 2010; Menting, 2013). This increased separation is only possible because the inter-subunit disulfides that link the  $\alpha$ CTs were deleted in the crystallized fragment of IR. In an intact receptor, the altered position of  $\alpha$ CT following binding of one ligand would restrict the reach of the opposite  $\alpha$ CT, likely hindering assembly of the second binding site and contributing to negative cooperativity.

The intrinsic propensity of the TMs to associate may also contribute to negative cooperativity. Disruption of the L1:Fn2'-3' interaction by ligand binding removes a restraint on the relative position of the TMs that served as a barrier to TM association (Figure 1.5). Decreasing this barrier may shift the energetic balance between maintenance of the L1':Fn2-3 interaction at the unliganded site and the intrinsic propensity of the TMs to associate and couple TM association to rearrangement of elements of the unliganded binding site. This mechanism may underlie the restoration of negative cooperativity of soluble IR ECDs when their C-termini are artificially tethered (Hoyne, 2000).

The results discussed so far implicate ECD-enforced TM separation as a key feature maintaining IGF1R and IR inactive states. The question then arises of how TM separation maintains the kinases in an inactive state. The simplest explanation is that spatial separation is sufficient to preclude one kinase from phosphorylating the other. The N-termini of subunits of an IGF1R kinase dimer captured in an apparent trans phosphorylation state are  $\sim 50$  Å apart (Wu, 2008), however, and the 27 residues separating each kinase domain from its TM could easily span the distance needed to reach the  $\sim 120$  Å separation imposed by the ECD. The IGF1R intracellular juxtamembrane regions must thus adopt some structure in the inactive state to preclude autophosphorylation, but secondary structure prediction algorithms only identify a short 7 amino-acid  $\beta$  strand in this region. Outside of three positively charged residues immediately following the TM, no conserved regions of positive charge comparable to the region in EGFR thought to mediate interactions with the membrane are present (Jura, 2009). Conversely, the fourfold enhancement of kinase activity when the juxtamembrane region is present implies some structure in this region in the active state. Attempts to

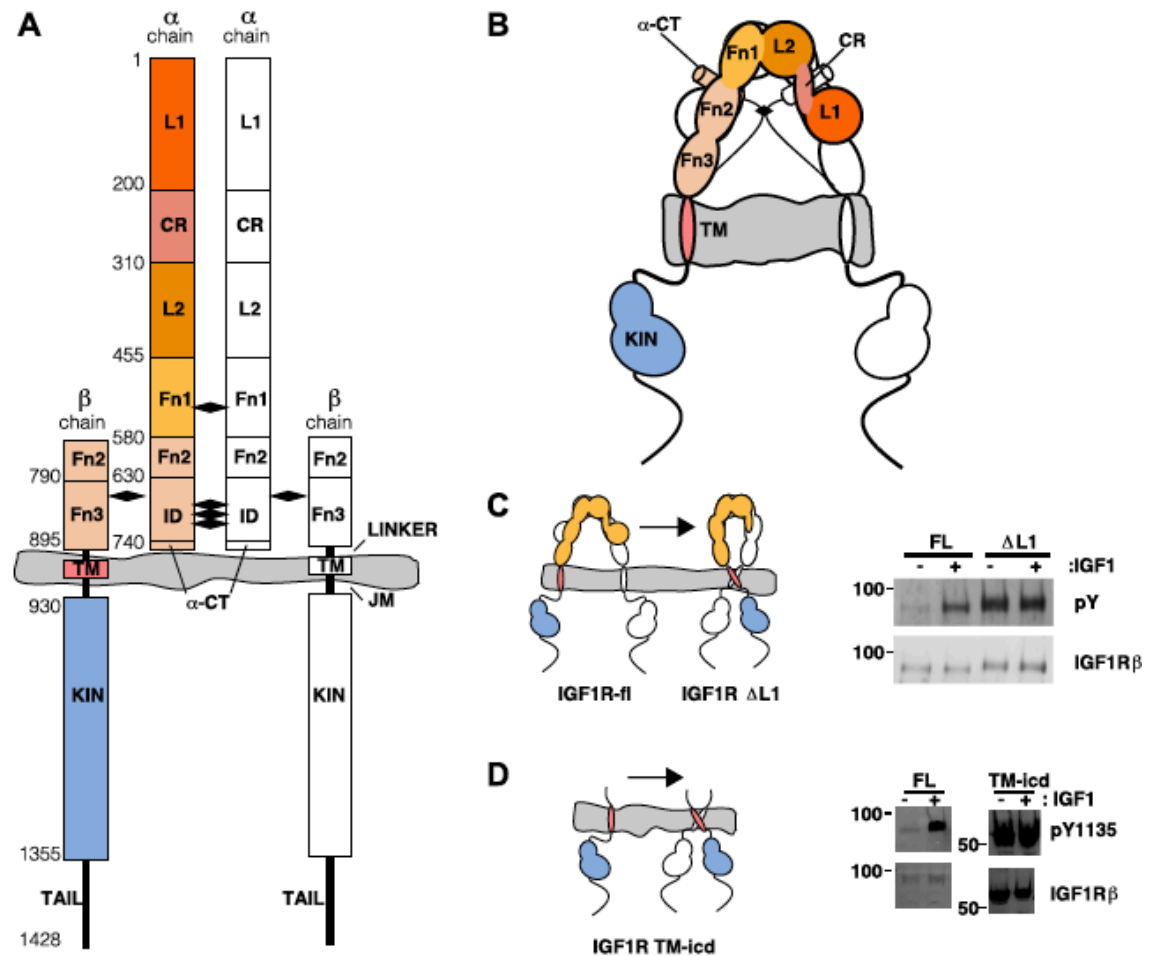
identify regions of the IGF1R kinase important for maintaining the inactive state by either mutagenesis or inspection of kinase crystal lattices for a repeated interaction proved unsuccessful, however. No IGF1R kinase surface mutations led to higher basal activity, and no repeated interactions or regions of positive charge are apparent in IGF1R kinase crystals that could mediate self interactions or interactions with the plasma membrane as has been proposed for EGFR (Arkhipov, 2013).

In contrast, the absence of allosteric enhancement of the activity of phosphorylated IGF1R implies the absence of a stable interaction between kinase domains in the active state. Consistent with this conclusion, surface mutagenesis of the IGF1R kinase demonstrates the absence of an EGFR-like asymmetric kinase dimer. Absence of allostery in the active state may be a general feature of RTKs whose activity is enhanced by multiple activation loop phosphorylations, including FGFR, Met, VEGFR, Alk, MUSK, Kit, and their homologs (Lemmon and Schlessinger, 2010). Transient intermediate states may direct the order of tyrosines phosphorylated, as observed for FGFR (Lew, 2009), but once fully phosphorylated we suspect the activity of these kinases will also prove independent of allosteric stimulation. This behavior contrasts with that of EGFR, which relies on allostery via a kinase asymmetric dimer for activation rather than phosphorylation of the activation loop (Zhang, 2006).

Although IGF1R ICDs do not appear to form a stable interaction in the active state, TM dimerization does appear to occur upon ligand binding. This observation is consistent with previous studies demonstrating a coupling between ligand binding and association of IR ECD C-termini as well as results with chimeric receptors. When the IR ECD is expressed in soluble form it binds with nearly 1000-fold lower affinity than full-

length IR, but when the C-termini of the ECD are fused to either a leucine zipper (Hoyne, 2000) or an immunoglobulin Fc region (Bass, 1996), high-affinity, native-like ligand binding is restored (Schaffer, 1994). TM dimerization is thus not only consistent with high-affinity ligand binding but promotes it. Furthermore, a chimeric receptor with an IR ECD and an EGFR TM and ICD can be activated by Insulin (Riedel, 1986). As EGFR activation involves TM dimerization (Endres, 2013), the Insulin-bound IR ECD is also compatible with TM dimerization.

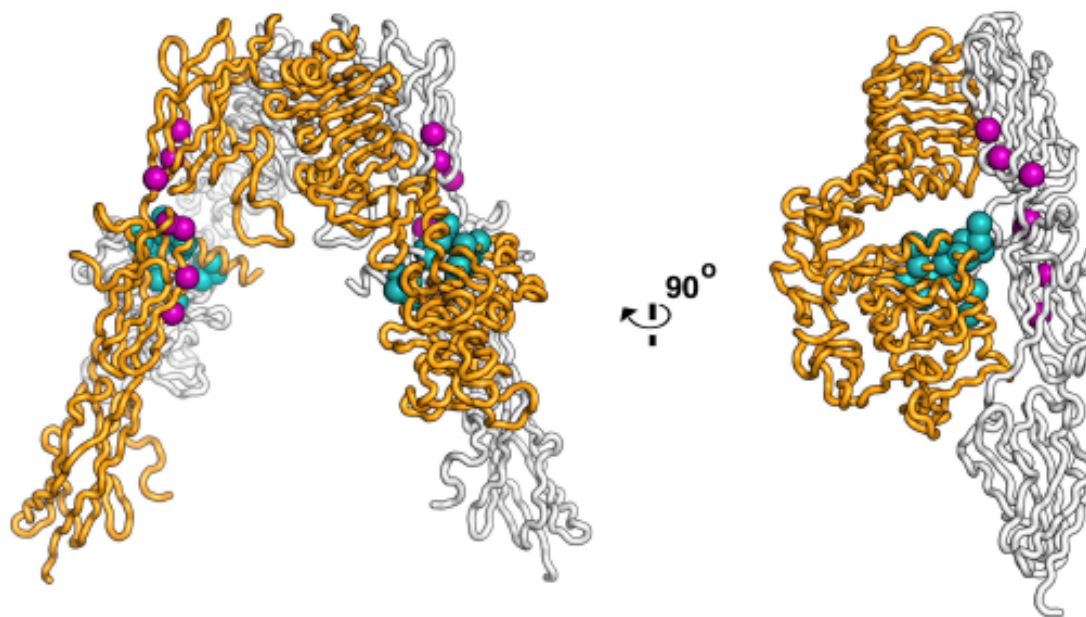
We present here a new molecular model for regulation of IGF1R and IR activity. The key feature of this model is maintenance of an inactive state by ECD-enforced separation of the TMs in the absence of ligand. Ligand binding releases this constraint, resulting in TM association and unleashing of the intrinsic propensity of the kinase regions to autophosphorylate and activate. This model is consistent with recent crystal structures of the IR ECD and suggests that some structure in the intracellular region precludes autophosphorylation in the unliganded state of the receptor. Conversely, enzymatic studies show that once phosphorylated and active, no structure outside of the kinase and juxtamembrane regions is needed to stimulate full kinase activity. This model provides a simple molecular context for understanding several features of IR/IGF1R activation and suggests directions for future study. It also suggests that chimeric proteins with IR or IGF1R ECDs may prove valuable for assessing the role of TM separation in other signaling systems or as ligand-dependent activity switches when fused, for example, to split enzymes.



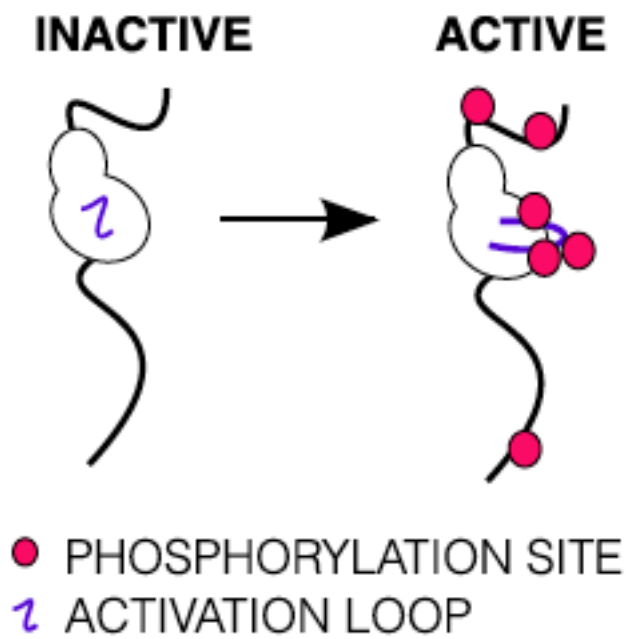
**Figure 1.1. The extracellular domain of IGF1R autoinhibits IGF1R activity. (A, B)** Schematic and cartoon representations of IGF1R. Numbering refers to human IGF1R excluding the signal sequence. Disulfide bonds are indicated by black diamonds. One  $\alpha\beta$  subunit is shown in outline and the other in color. (C) Cartoon of IGF1R full-length (IGF1R-fl) or an IGF1R fragment lacking the L1 domain (IGF1R  $\Delta$ L1) (left). HEK293 cells expressing either IGF1R-fl (FL) or IGF1R  $\Delta$ L1 ( $\Delta$ L1) were incubated with or without IGF1. IGF1R proteins were immunoprecipitated from normalized cell lysates and



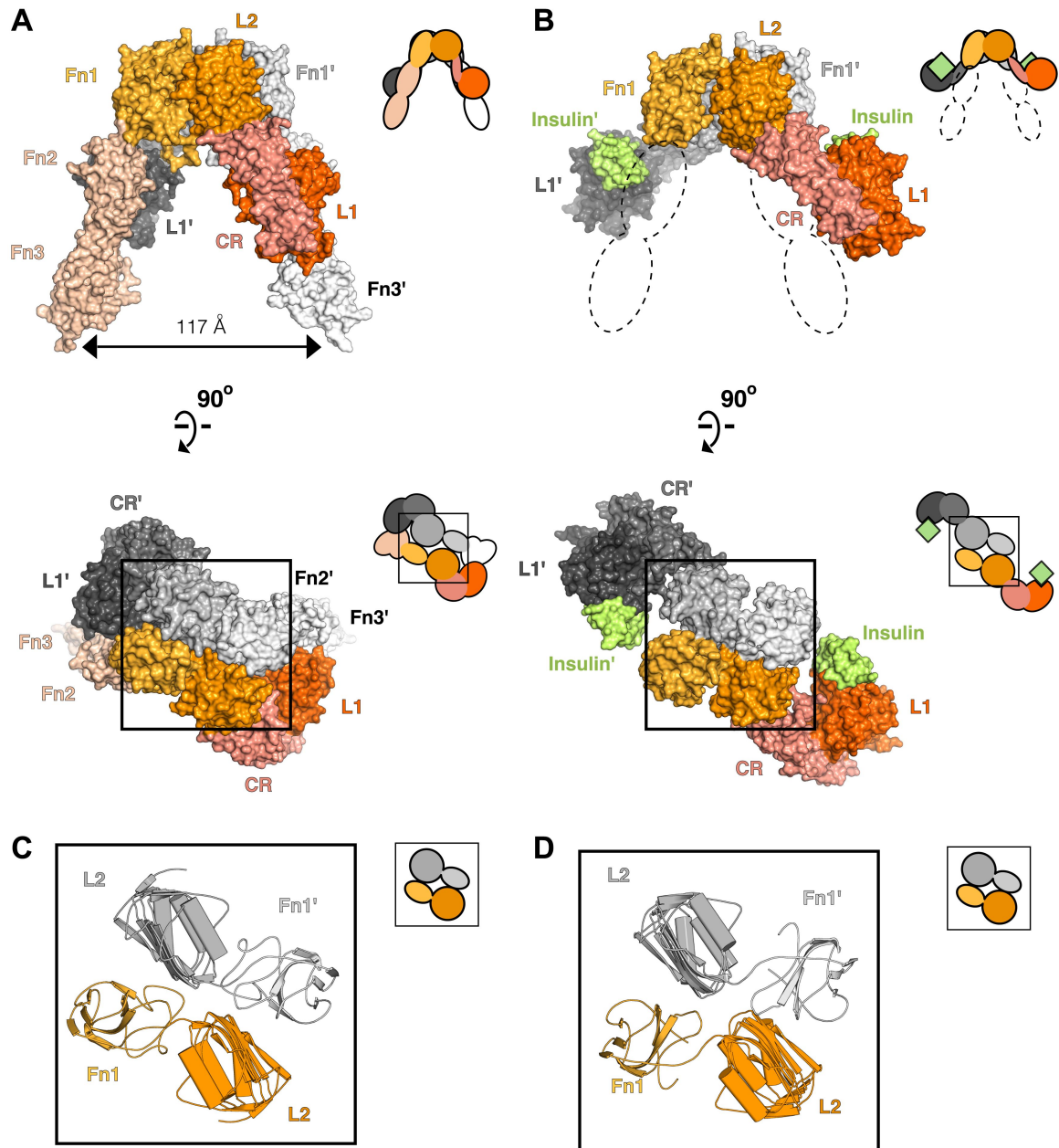
levels of autophosphorylation or IGF1R expression detected by Western blot with anti-phosphotyrosine (pY) or anti-IGF1R $\beta$  antibodies (right). Molecular weight standards indicated to the left of each panel (D). Diagram of IGF1R variant lacking the ECD (IGF1R TM-icd) (left). Western blots probed with anti-pY1135, which recognizes a phosphotyrosine on the activation loop of the IGF1R kinase domain, and anti-IGF1R $\beta$  of normalized lysates from cells expressing IGF1R-fl (FL) or IGF1R TM-icd (TM-icd), in the presence and absence of IGF1 (right). All panels are from the same blot.



**Figure 1.2 Residues on IR important for ligand binding.** Worm diagram of the unliganded IR ECD with one  $\alpha\beta$  subunit colored white and the other orange. Residues implicated in forming the Site 1 binding surface are shown as cyan spheres and residues implicated in forming the Site 2 binding surface are shown in purple spheres (Mynarcik, 1996; Whittaker, 2001; Whittaker and Whittaker, 2005; Whittaker, 2008).

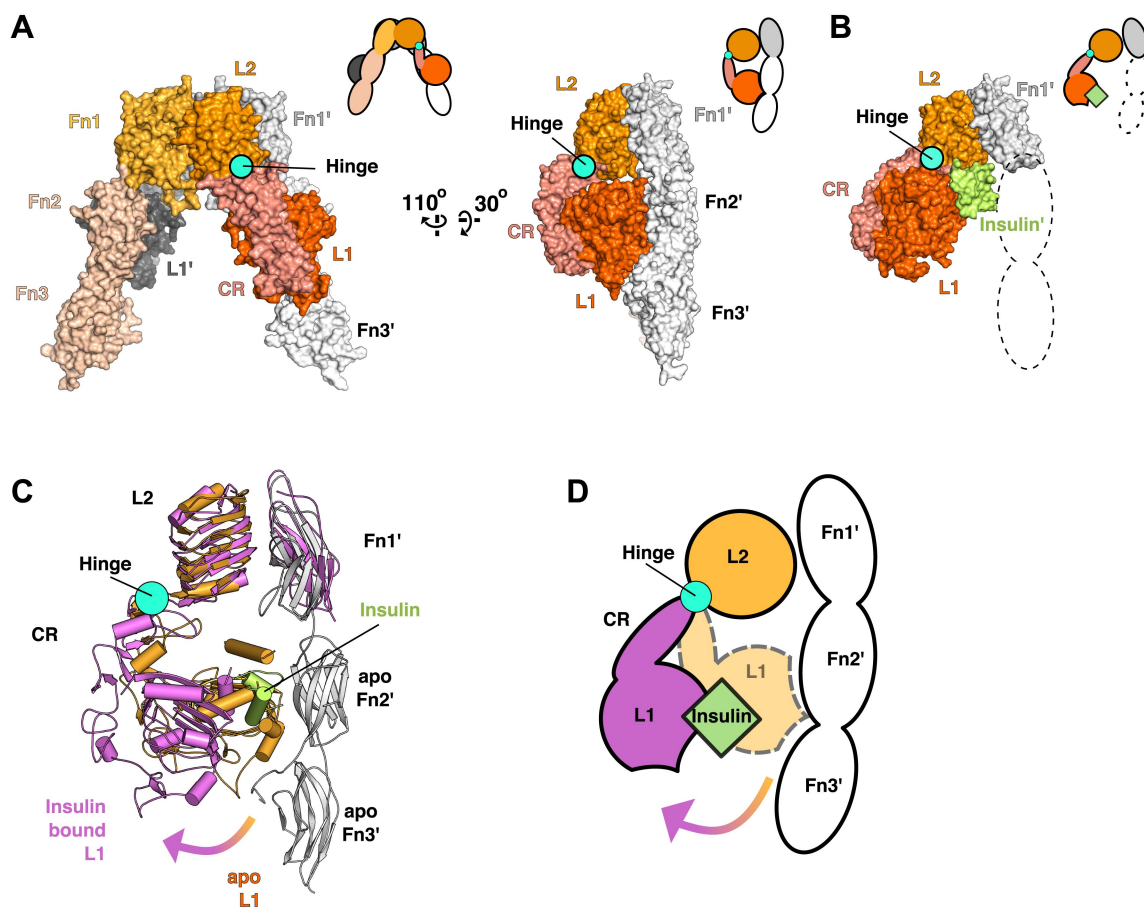


**Figure 1.3 Phosphorylation of the ICD.** Cartoon of IGF1R-icd in the inactive or active, phosphorylated conformations. The activation loop is indicated by a purple line and phosphorylation sites as pink circles.



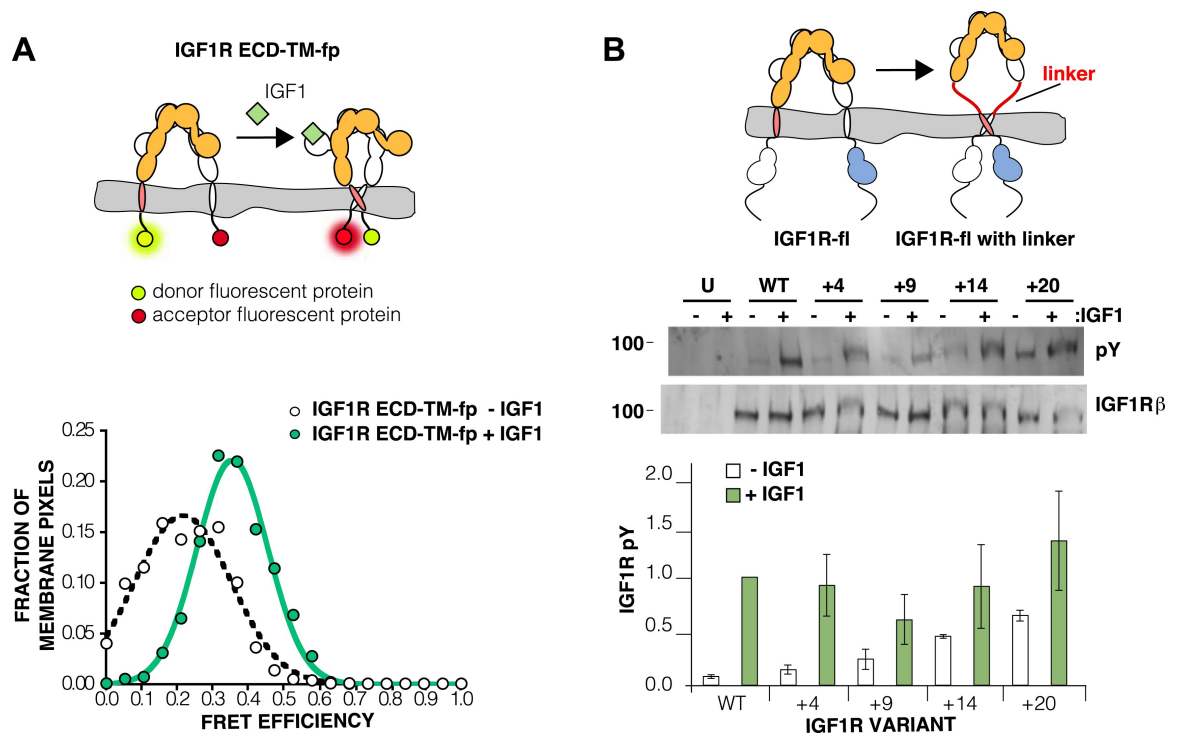
**Figure 1.4 Conserved IR ECD dimer interaction.** Surface representations of (A) the unliganded IR ECD dimer (PDB 3LOH; Smith, 2010) and (B) the IR ECD fragment bound to Insulin (PDB 3W14; Menting, 2013). Domains of one  $\alpha\beta$  subunit are colored in shades of orange, the other in shades of gray, and Insulin in light green. N-terminal domains are in darker shades. The conserved L2-Fn1:L2'-Fn1' interface is marked by a black box. Ribbon diagrams of these domain pairs are shown for the unliganded (C) and

liganded (D) IR ECD dimer. A cartoon of domains shown are included at the upper right of each structure. Dashed lines represent the missing Fn2-3 domain tandems.



**Figure 1.5 Ligand-induced conformational change in the IR ECD.**(A) Surface representation of the unliganded IR ECD (PDB 3LOH; Smith, 2010) oriented to show the L1:Fn2'-3' interface. The hinge between CR and L2 is indicated by a cyan circle. Cartoons of domains shown are to the upper right of each structure. (B) Surface representation of ligand-bound IR ECD fragment (PDB 3W14; Menting, 2013) in the same orientation as in A. The missing Fn2-3 domains are indicated by dashed lines. (C) Superposition of the ligand-bound IR structure (purple ribbon) onto the unliganded structure (orange and white ribbon) using the L2-Fn1 domain pairs. Insulin is displayed in light green. An arrow shows the direction of the relative movement of unsuperposed L1-CR domains about a hinge (cyan circle) between the CR and L2 domains.

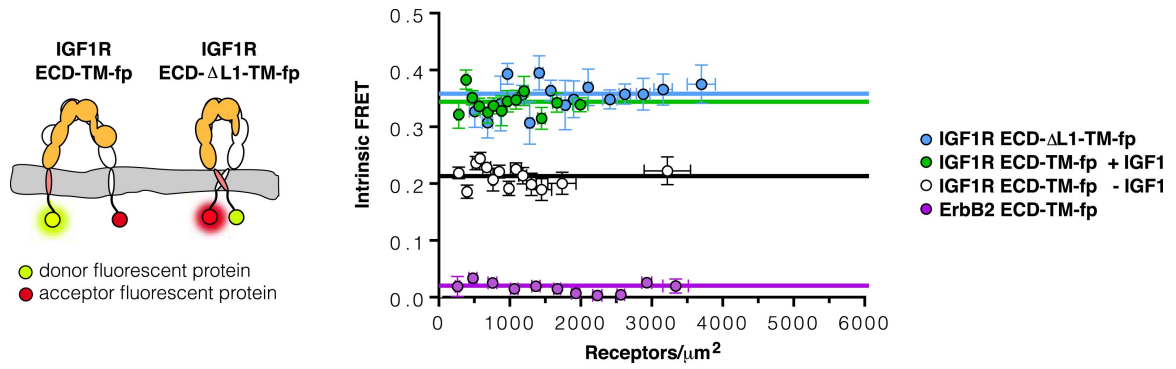
Complementary halves of each receptor subunit were omitted for clarity. (D) A cartoon showing the movement of the L1-CR domain tandems.



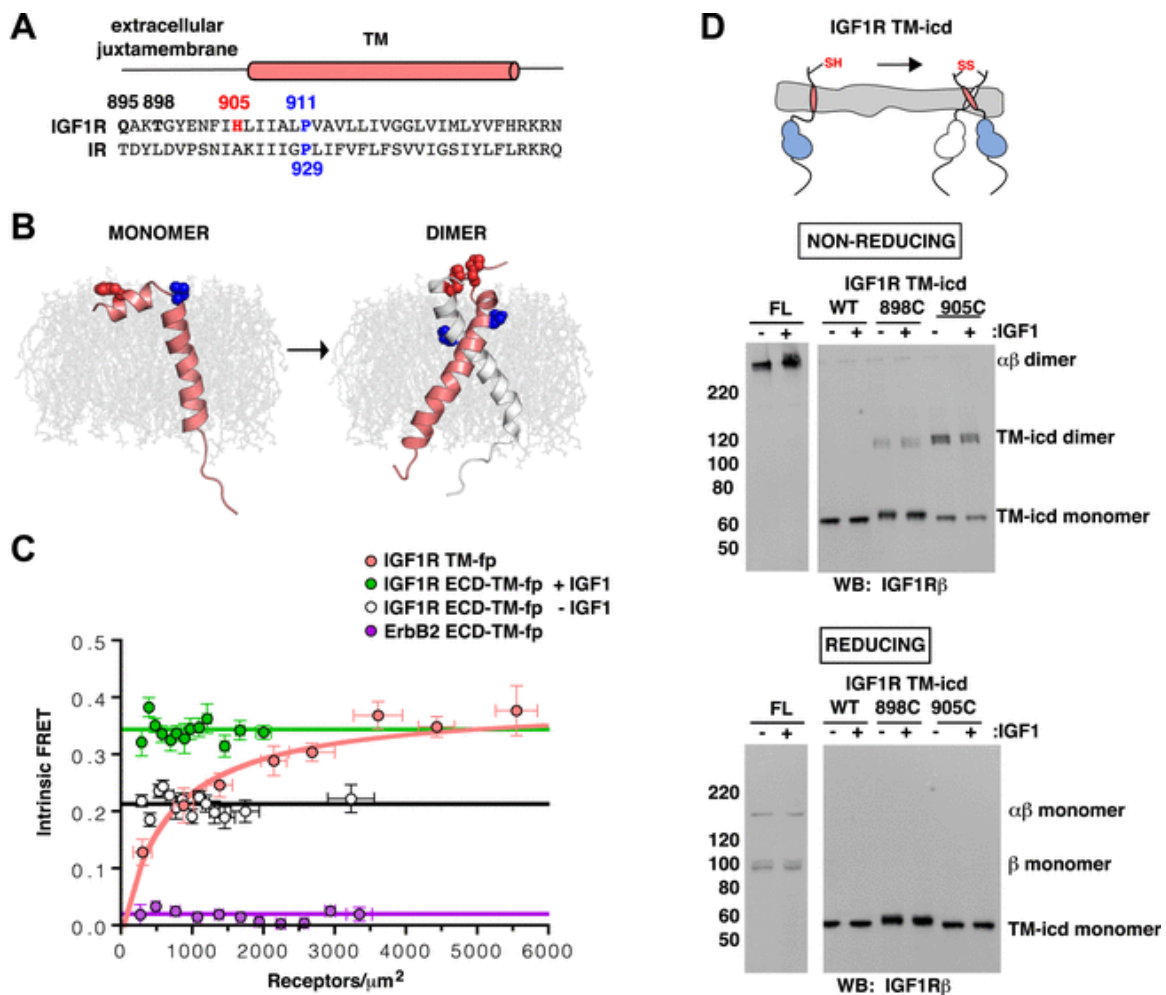
**Figure 1.6 The IGF1R ECD maintains TM separation.** (A) Cartoon of the IGF1R ECD-TM-fp variant utilized in the live cell FRET assay (top). Frequency distribution of FRET efficiency values per membrane pixel of IGF1R ECD-TM-fp in the presence (green) or absence (white) of IGF1 (bottom). Data were fit to a normal distribution with peak values of  $0.34 \pm 0.09$  and  $0.21 \pm 0.13$  in the presence and absence of IGF1, respectively. (B) Cartoon of the IGF1R-fl and IGF1R variants with flexible linkers (top). Representative western blots probed with either anti-pY and anti-IGF1R $\beta$  antibodies of immunoprecipitated, normalized cell lysates of untransfected cells (U) or cells transfected with IGF1R-fl (WT) or IGF1R variants with an additional 4 (+4), 9 (+9), 14 (+14), or 20 (+20) glycine and serine residues per  $\alpha\beta$  chain (middle). Bar graph of average IGF1R phosphorylation normalized to total receptor concentration ( $\pm$ s.e.m.) from at least three separate experiments, except for +14 which was from 2. Results from cells incubated in



the absence of IGF1 are shown in white and in the presence of IGF1 shown in green  
(bottom).

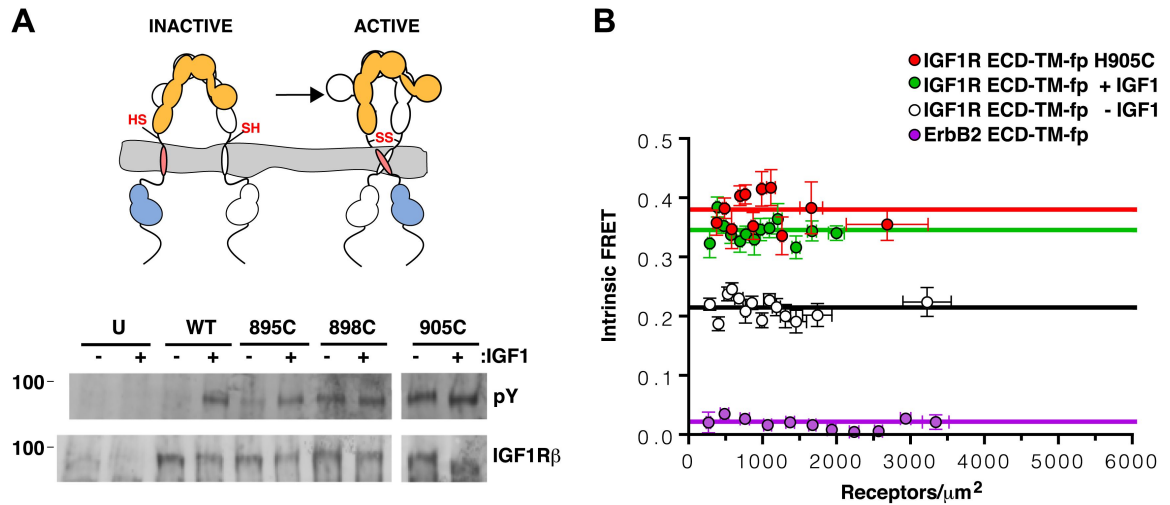


**Figure 1.7 The TMs associate in IGF1R ECD-ΔL1-TM-fp.** Cartoons of IGF1R ECD-TM-fp and IGF1R ECD-ΔL1-TM-fp used in FRET assays (left). The intrinsic FRET of IGF1R ECD-ΔL1-TM-fp (blue) in vesicles plotted as a function of receptor concentration and fit to a horizontal line with a Y intercept of  $0.36 \pm 0.02$  (right). Each data point represents the binned average of at least eight vesicles with error bars ( $\pm$ s.e.m. in both x and y). Shown for reference are the data and fits of IGF1R ECD-TM-fp without IGF1 (white), with IGF1 (green), or the monomeric control ErbB2 ECD-TM-fp (purple).

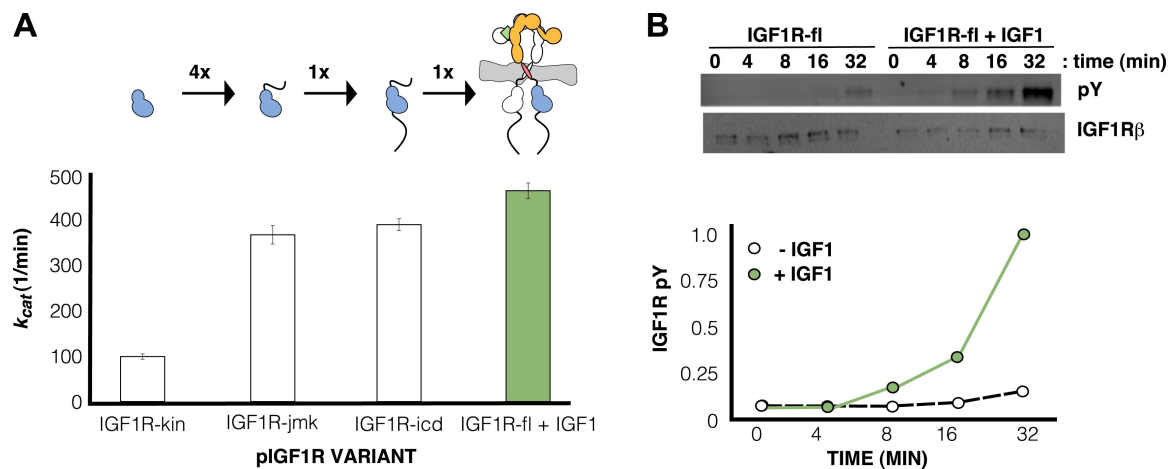


**Figure 1.8 IGF1R TMs associate.** (A) Sequence alignment of human IGF1R and IR extracellular juxtamembrane and TM regions. The position of the TM is indicated above the alignment. Bold lettering highlights residues targeted for substitution. H905 (red) and P911 (blue) are colored. (B) IGF1R TM monomer and dimer structures observed in MD simulations are shown in ribbon (white and pink). Shown in spheres are P911 (blue) and H905 (red). The lipid bilayer is in light gray. (C) The intrinsic FRET efficiency for IGF1R TM-fp plotted (pink) as a function of TM concentration. Also shown for reference are a monomeric control (ErbB2 ECD-TM-fp) in purple and the IGF1R ECD-TM-fp in the presence (green) or absence (white) of ligand. Each data point represents the binned average of at least eight vesicles ( $\pm$ s.e.m. in both x and y). For IGF1R TM-fp, the data

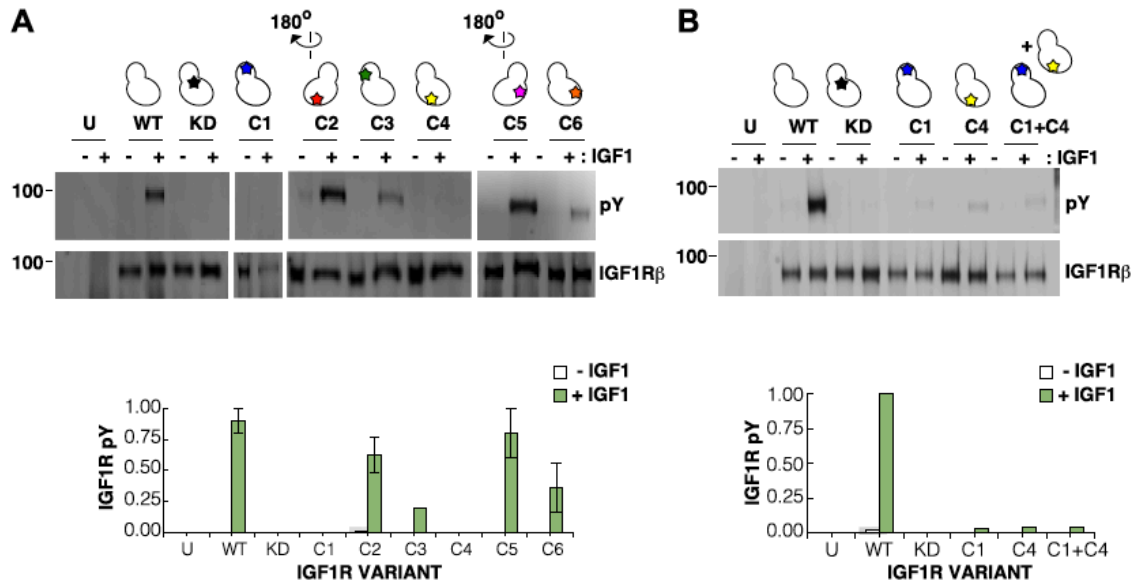
were fit to a two state association model with a peak value of  $0.38 \pm 0.02$ . (D) Cartoon of IGF1R TM-icd with the positions of single-site cysteine mutations indicated in red (top). Anti-IGF1R $\beta$  Western blots of lysates of HEK293 cells expressing either IGF1R-fl or IGF1R TM-icd fragments (WT, 898C or 905C) incubated with or without IGF1 and analyzed on non-reducing (middle) or reducing gels (bottom).



**Figure 1.9 TMs associate in active IGF1R-fl.** (A) Cartoon of IGF1R-fl with cysteine substitutions (top). Western blots of immunoprecipitated normalized cell-lysates from untransfected cells (U) or cells transfected with IGF1R-fl (WT) or IGF1R-fl proteins with cysteine substitutions (895C, 898C, or 905C) incubated the presence or absence of IGF1 (bottom). All panels are from the same blot. (B) Plot of the intrinsic FRET efficiency of IGF1R ECD-TM-fp with the H905C substitution (red) and fit to a horizontal line with a value of  $0.38 \pm 0.02$ . Each data point represents the binned average of at least eight vesicles ( $\pm$ s.e.m. in both x and y). For reference, the intrinsic FRET efficiencies and fits for IGF1R ECD-TM-fp with ligand (green) or without ligand (white) and the ErbB2 ECD-TM-fp negative control (purple) are also shown

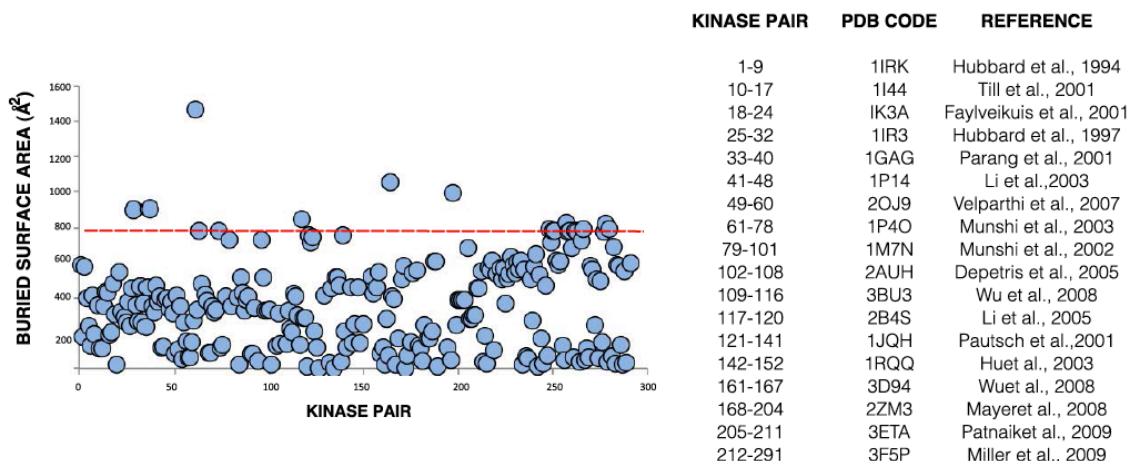


**Figure 1.10 IGF1 stimulates autophosphorylation.** (A) Bar graphs of  $k_{cat}$  ( $\text{min}^{-1}$ ) measurements are shown for phosphorylated forms of indicated IGF1R proteins. Diagrams of the specific fragments and the fold differences in activation are shown. (B) Anti-phosphotyrosine (pY) and anti-IGF1R $\beta$  Western blots of IGF1R-fl as a function of time in the presence and absence of IGF1 (top). Quantification of band intensities in the presence (green circles) and absence (white circles) of IGF1 (bottom).



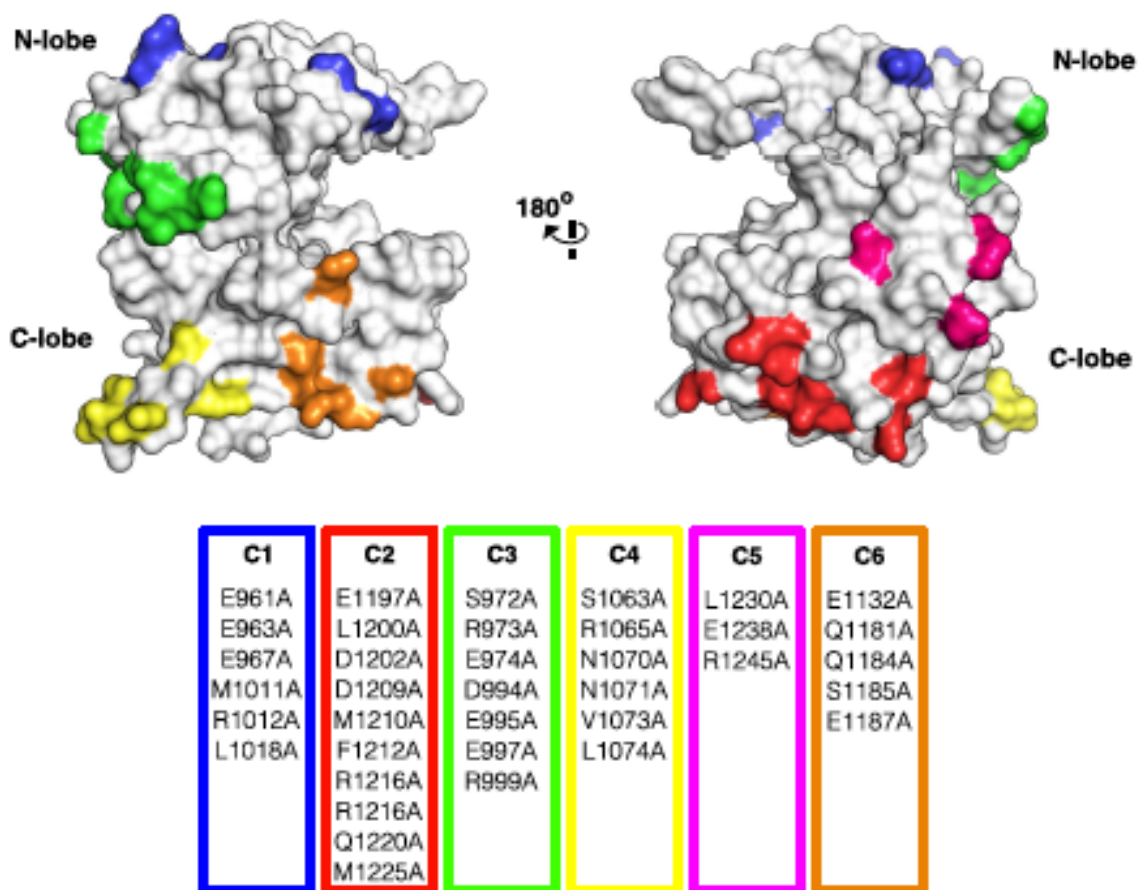
**Figure 1.11 Surface mutations on the IGF1R kinase domain disrupt receptor**

**activation.** (A) Anti-pY and anti-IGF1R $\beta$  Western blots of tagged IGF1R proteins immunoprecipitated from cells incubated in the presence or absence of IGF1. HEK293 cells were transiently transfected with IGF1R-fl (WT), IGF1R bearing the kinase-inactivating mutation D1105A (KD), IGF1R with clusters of surface mutations (C1-C6), or not transfected (U). Approximate locations of the mutated clusters indicated by a star on the kinase domain cartoons above each lane (top). Bar graphs showing mean receptor phosphorylation ( $\pm$ s.e.m.) from three independent experiments normalized to total receptor expression (for C3, only one experiment). White bars indicate cells incubated without IGF1 and green bars cells with IGF1 (bottom). (B) Western blots of IGF1R phosphorylation when C1 and C4 variants are co-transfected (top). Bar graph of receptor phosphorylation normalized to total receptor expression with the same coloring as in (A) (bottom).

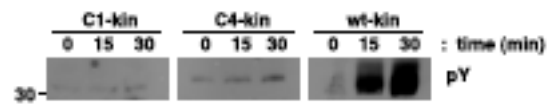


**Figure 1.12 Buried surface analysis of crystal lattice pairs of IR and IGF1R kinase domains.** A scatter plot of the buried surface area ( $\text{\AA}^2$ ) of each crystal lattice pair of IR or IGF1R kinase domains. A dashed line (red) at  $800 \text{ \AA}^2$  represents the minimal threshold of a biological interface (Bahadur, 2004) and a list of structures used in analysis and the numbering of the corresponding kinase pairs generated for each (Hubbard, 1994; Hubbard, 1997; Favelyukis, 2001; Parang, 2001; Pautsch, 2001; Till, 2001; Munshi, 2002; Hu, 2003; Li, 2003; Munshi, 2003; Depetris, 2005; Li, 2005; Velaparthi, 2007; Mayer, 2008; Wu, 2008; Miller, 2009; Patnaik, 2009).



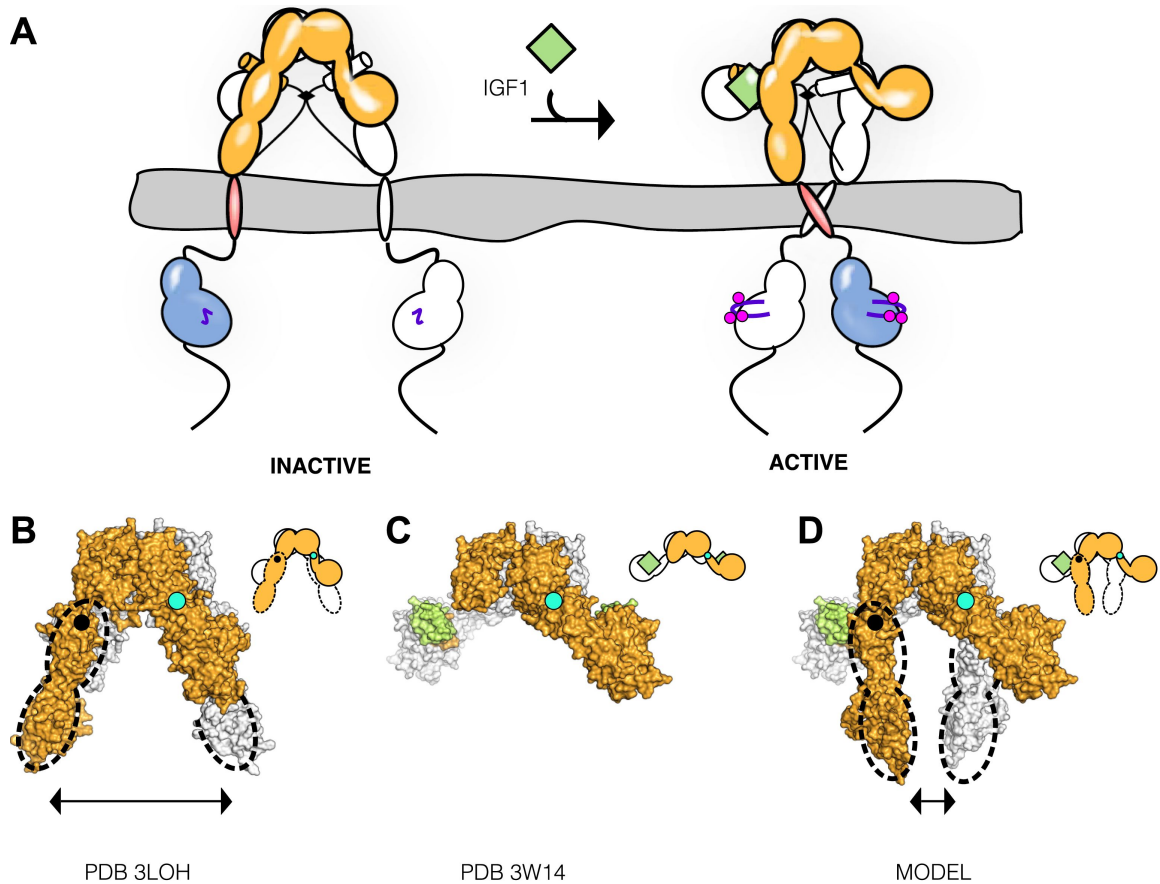


**Figure 1.13: Surface analysis of IGF1R kinase domain.** The IGF1R kinase displayed in surface representation showing the location of each cluster colored according to the table of alanine substitutions made for each cluster (PDB 1P4O, Munshi, 2002).



**Figure 1.14 Analysis of the autophosphorylation of IGF1R kinase clusters.**

Autophosphorylation of IGF1R-kin wild-type (WT), cluster 1 (C1) and cluster 4 (C4) as a function of time monitored by anti-pY Western-blot. All lanes shown are from the same blot.



**Figure 1.15 Model of IGF1R activation.** (A) Cartoon model of IR family activation. The IDs are shown as black lines in the ECDs, the disulfide linkages as black diamonds, phosphorylation as pink circles, activation loops as purple lines, ligand as green diamonds, and  $\alpha$ CTs as cylinders. Surface representations and corresponding cartoon diagrams of (B) the ECD of unliganded-IR (PDB 3LOH; Smith, 2010), (C) insulin bound to the IR ECD fragment (PDB 3W14; Menting, 2013), or (D) a model of the entire IR ECD bound to two Insulin molecules. For each panel, one  $\alpha\beta$  subunit is colored white and the other orange, and Insulin green. Dashed lines outline the Fn2-3 domain tandems. The hinge point between F1 and Fn2-3 domains is indicated by black circles and cyan circles indicate the hinge point between CR and L2 domains.

## **Chapter 2. General and Specific Receptor tyrosine Kinase Signaling Mechanisms Revealed by Analysis of Chimeric EGFR and IR Family Receptors**

### **Abstract**

Ligand-induced dimerization of receptor tyrosine kinases (RTKs) is a generally accepted mechanism of activation for all classes of RTKs except the Insulin Receptor (IR) family. In contrast to all other RTK families, IR family members exist as preformed disulfide linked dimers on the cell surface. Recent evidence suggests that the ectodomain (ECD) of IR, and that of its homolog the type 1 insulin-like growth factor receptor (IGF1R), maintains separation of subunit transmembrane regions in the absence of ligand. Ligand binding releases this ectodomain autoinhibition and allows the transmembrane domains to come together and the intracellular kinase domains to transphosphorylate. Parallels between the IR/IGF1R mechanism of activation and ligand-induced dimerization suggested that the IR/IGF1R ECDs may be able to regulate other families of RTKs, and we set out to understand how the IR/IGF1R activation mechanism relates to EGFR activation. We find that the IR/IGF1R ECD fails to activate the kinase domain of EGFR in either IR-EGFR or IGF1R-EGFR chimeric receptors. In contrast, the EGFR ECD is able to activate the IR kinase domain in an EGFR-IR fusion in an EGF-dependent fashion. The IR and EGFR mechanisms of activation thus appear to share some but not all features.

### **Introduction**

Receptor tyrosine kinases (RTKs) are a class of cell surface receptors defined by an extracellular domain (ECD), a single-pass transmembrane (TM) domain, and an intracellular region containing a tyrosine kinase domain (Reviewed in Lemmon and Schlessinger, 2010). These receptors respond to various extracellular growth factor signals with enhanced intracellular kinase activity that in turn activates downstream signaling pathways involved in cell growth and differentiation. There are fifty-eight RTKs in humans that assort into twenty classes based on homology. Both topology and common mechanisms define the RTK family. It is generally accepted that RTKs exist on the cell surface predominantly as monomers, likely in equilibrium with a small population of inactive preformed dimers, and binding of extracellular ligand induces receptor dimerization and activation. Increasingly sophisticated methods of analyzing receptor structure and ligand binding suggest complex levels of regulation of RTK activity that complicate even the well-accepted notions of ligand-induced dimerization.

Perhaps the most well studied, and often considered canonical, RTK is the Epidermal Growth Factor Receptor (EGFR) (Lemmon, 2014). EGFR is one of four members of RTK Class I, which also includes HER2, HER3, and HER4. Biochemical and structural studies of EGFR and other class I members are well advanced and thought in many instances to reflect general features of RTK behavior. In the absence of ligand, EGFR is predominately a monomer and adopts a tethered conformation that blocks the EGFR dimerization interface (Figure 2.1). The binding of EGF results in a conformational change whereby the dimerization interface is exposed, promoting receptor dimerization and kinase domain transphosphorylation. Unique to EGFR is the formation of an intracellular, asymmetric dimer between kinases where the C-lobe of one

kinase domain contacts the N-lobe of the second kinase domain in the dimer and allosterically stimulates its kinase activity (Zhang, 2006). Complicating the simplistic notion of an EGFR monomer dimerizing in the presence of ligand is experimental evidence for preformed, inactive EGFR dimers (Endres, 2013). The nature of preformed dimers is not well understood although homologs of EGFR have begun to provide insight. The *Drosophila* EGFR ectodomain has been crystallized as a dimer in both the presence and absence of ligand, revealing a distinct ECD dimer in the unliganded state (Alvarado, 2009; Alvarado, 2010). More recently, a soluble construct of the ECD of Let-23, the *C. elegans* EGFR homolog, was determined to be dimeric in solution (Freed, 2015). In cells, however, Let-23 maintains its kinase domain in an inactive state until the cognate ligand is added. How ligand binding to a preformed dimer of human EGFR, or dimeric EGFR homolog, activates the receptor must include a conformational change whose nature is not well understood.

Further complicating the notion of ligand-induced dimerization is RTK Class II, the Insulin Receptor (IR) family, which in addition to IR includes the type 1 insulin-like growth factor receptor (IGF1R) and the insulin related receptor (IRR). In contrast to all other RTK classes, IR, IGF1R, and IRR exist on the cell surface as preformed, disulfide-linked dimers. Ligand-induced dimerization cannot explain how IR is activated upon ligand binding. Recent structural studies of IR have elucidated both snapshots of apo and ligand bound forms of portions of the IR extracellular domain (McKern, 2006; Menting, 2013). Furthermore, biochemical and biophysical studies of IGF1R have suggested that the role of the IGF1R ECD is autoinhibitory and serves to maintain transmembrane region separation in the absence of ligand (Kavran, 2014). Ligand binding releases a

constraint in the ECD and allows the transmembrane regions to move closer together, resulting in kinase domain activation.

This simple notion of transmembrane regions apart, receptor-off and transmembrane regions close together, receptor-on draws immediate parallels to ligand-induced dimerization. In the case of IR and IGF1R, ligand binding allows the transmembrane and intracellular regions to come close together and dimerize and the kinase domains to autophosphorylate. Comparisons, then, between the mechanism of activation of IR and IGF1R with other RTK classes governed by ligand induced dimerization appeared worth investigating.

The validity of this comparison is strengthened by widely accepted experiments involving chimeric IR and EGFR receptors. Chimeric receptors with either the IR ECD fused to the EGFR TM and intracellular domain (ICD) or the EGFR ECD and TM domain fused to the IR ICD were activated by the cognate ligand of the ECD (Riedel, 1986; Riedel, 1989). These experiments were performed in the early stages of RTK characterization, and the concept that these chimeric receptors are active has influenced mechanistic studies of RTKs for multiple decades.

Given recent advances in our knowledge of how IGF1 and insulin activate their receptors, we set out to replicate the IR-EGFR chimeric receptor experiment and extend the chimeric receptor approach to other RTK families. Does the IR mechanism of activation extend to EGFR, as previously reported, and other RTK family members? Is the separation of transmembrane regions by IR sufficient to keep the EGFR kinase domains inactive? Is ligand-induced transmembrane association by the IR/IGF1R

ectodomains sufficient to activate the EGFR kinase? Is the preformed dimer of IR really mechanistically distinct from the 19 other monomeric RTKs?

## **Materials and Methods**

### *Buffers*

Homemade RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% (v/v) NP40, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulfate was stored at 4°C for up to six months. Phosphate buffered saline (PBS; 0.138 M sodium chloride, 0.0027 M potassium chloride, pH 7.4). PBS-T (PBS plus 1% (v/v) Tween 20).

### *Reagents*

PureLink<sup>®</sup> HiPure Plasmid Midiprep Kit (Invitrogen by Life Technologies, K2100-05). Polyethylenimine, PEI, linear MW 25,000 (Polysciences, Inc., 23966-2). Pierce<sup>™</sup> BCA Protein Assay (Life Technologies, 23225). Protein G Sepharose 4 Fast Flow (GE Healthcare, 17-0618-01). Dynabeads<sup>®</sup> Protein G (Novex by Life Technologies, 1003D). Novex 4-20% Tris-Glycine Mini Protein Gels, 1.0mm, 15 well (Life Technologies, EC6025). iBlot<sup>®</sup> Dry Blotting System and iBlot<sup>®</sup> Transfer Stacks, Nitrocellulose (Life Technologies, IB301001). Thermo Scientific<sup>™</sup> Pierce<sup>™</sup> ECL 2 Western Blotting Substrate (Fischer Scientific, 80197).

Western Blots were placed on Apollo Plain Paper Copier Transparency Film (VPP100c) to image with the Typhoon<sup>™</sup> FLA 9500 biomolecular imager (GE Healthcare Life Sciences). Typhoon settings for ECL 2 were used(laser: 473 nm (Blue LD laser), filter: BPB1 (530DF20), PMT 400, and 100 µm resolution.

### *Western Blot Protocol and Antibody Solutions*



Nitrocellulose membranes were blocked for 30 minutes in blocking solution, PBS-T plus 3.3% (w/v) BSA prior to primary antibody. Unless otherwise indicated, antibodies were diluted in blocking solution and membrane incubation with diluted antibodies proceeded for 1 hour at room temperature with gentle agitation. After each antibody incubation, membranes were washed three times with PBS-T. Working dilutions were kept for up to two weeks at 4°C.

Antibodies are listed with name, clone identification number where applicable, vendor, catalog number and working dilution. To detect pan phosphotyrosine, primary antibody used was anti-Phosphotyrosine, clone 4G10<sup>®</sup> (EMD Millipore, 05-321, 1:1000) and secondary antibody used was goat anti-mouse IgG<sub>2b</sub>-HRP (Santa Cruz Biotechnology, sc-2062, 1:3000). To detect HA epitope tag, primary antibody was anti-HA High Affinity (Roche, 11 867 423 001, 1:1000) and secondary antibody was goat anti-rat IgG-HRP (Santa Cruz Biotechnology, sc-2006, 1:2000). To detect Myc epitope tag, primary antibody was c-Myc, 9E10 (Santa Cruz Biotechnology, sc-40) and secondary antibody was goat anti-mouse IgG-HRP (Santa Cruz Biotechnology, sc-2006, 1:1000).

#### *Expression vector design*

Regions of cDNA encoding all RTK variants were cloned into modifications of the vector p $\alpha$ H. p $\alpha$ H is a vector derived from pHLSec (Aricescu, 2006) by substitution of an alternative multiple cloning site. Modifications to p $\alpha$ H included signal sequence removal, addition of a common set of epitope tags, and addition of a panel of sequences of the IGF1R, EGFR and IR ECDs to facilitate easy cloning of chimeric receptors. Vector descriptions can be found in Table 2.1.

Full length human IGF-1R (NM\_000875.4), IR, (NM\_000208.2) and EGFR (NM\_005228.3) cDNAs were used to generate full length and chimeric receptors. The following chimeric receptors were created to best reproduce previous IR/EGFR chimeras (Riedel, 1986; Riedel, 1989) and are also described in Table 2.2. IG1R-EGFR is a fusion between the IGF1R ECD (amino acids 1-936) and the EGFR TM-ICD (643-1210). IR-EGFR is a fusion between the IR ECD (1-953) and the EGFR TM-ICD (643-1210). EGFR-IR is a fusion between the EGFR ECD-TM (1-668) and the IR ICD (979-1382). IGF1R-IR is a fusion between the IGF1R ECD-TM (1-929) and the IR ICD (979-1382). Where applicable, chimeric clones were spliced together at sites within 1-2 amino acids of the chimeric fusion protein previously assayed (Riedel, 1986; Riedel, 1989). In the Riedel work, the EGFR transmembrane region was used in all chimeric receptors that contained EGFR.

Additional IGF1R-RTK chimeric receptors studied utilized the IGF1R ECD or IGF1R ECD-TM fused to the TM-ICD or ICD or the RTK of interest. An exhaustive list of all RTK clones can be found in Table 2.2. Transmembrane boundaries for the RTK of interest were guided by domain annotation of mRNA and protein sequences from NCBI. In general, two amino acids were added prior to the transmembrane region for TM-ICD fusions or after the TM for ICD fusions.

#### *Transient transfection of HEK293 and CHO cells*

Transient transfections followed methods previously described (Kavran, 2014). Sterile, midiprep DNA was used for cell-based assays. In brief, CHO or HEK293 cells were plated in six-well plates at  $0.5 \times 10^6$  cells/well and transiently transfected with

expression plasmids using polyethylenimine at an optimized ratio of 1  $\mu$ g DNA: 3  $\mu$ g PEI. Transient transfection proceeded for 18 hours.

#### *Generation of stably-transfected CHO cell lines*

CHO stable cell lines were generated by transient co-transfection of the plasmid of interest with pCDNA3.1, a vector containing the Neomycin resistance gene.

Transfection complexes containing PEI (30 $\mu$ g), RTK plasmid of interest (10 $\mu$ g) and pCDNA3.1 (0.1 $\mu$ g) were added to CHO cells seeded at  $1 \times 10^6$  cells/well in a 150mm plate.

After 24 hours, media was removed and replaced with fresh media containing 1mg/ml G418. G418 selection proceeded for 7 days. After 7 days, media containing G418 was changed every 2 days until colony formation was visible. Colonies were isolated using cloning rings and split evenly between 2 wells of a 24 well plate. One well was used to propagate the colony and the second was used to harvest cells and assay for expression.

As nearly all clones picked expressed the construct of interest, the number of clones picked per construct was lowered from 16 for initial stables to 6. Three different clones per construct were cryopreserved to represent a range of expression levels.

#### *Cell-based activity assays*

Cell-based assays followed methods previously established to assess IGF1R activity (Kavran, 2014). For stable cell lines, cells were plated at least 2 hours prior to serum starvation. For transient transfections, serum starvation began 18 hours after transfection. Three hours of serum starvation was followed by 5 minute ligand stimulation at 37°C with 20 nM IGF1, 200 nM insulin, or 100 ng/ml EGF as indicated. Wells were washed with ice-cold phosphate buffered saline and then lysed for 30 min at 4°C in 250  $\mu$ l of RIPA buffer supplemented with 1 mM activated  $\text{Na}_3\text{VO}_4$ , 1 mM PMSF,

Benzonase nuclease (Sigma). Total protein concentration of clarified lysates was determined using the BCA assay and lysates were normalized to lowest total protein content using RIPA buffer. Immunoprecipitation was performed by adding normalized cell lysates to protein G dynabeads pre-equilibrated with anti-HA antibody. To pre-equilibrate beads, 200 ng of anti-HA was mixed with 0.6 mg dynabeads in PBS plus 0.01% Tween 20 by rotation for 30 minutes at room temperature. Immunoprecipitations proceeded by rotating samples overnight at 4°C. The following day, beads were washed three times with 1 ml of RIPA buffer supplemented with 1 mM activated  $\text{Na}_3\text{VO}_4$ . Immunoprecipitates were eluted by addition of 30  $\mu\text{l}$  of 2X SDS loading buffer containing 10% fresh  $\beta$ -mercaptoethanol to beads. Equal amounts of eluted proteins were resolved using 4-20% Tris-Glycine gels and protein was transferred to nitrocellulose. Blots were developed using both anti-pY and anti-myc or anti-HA. Anti-HA was used for detection only when anti-myc detection level was low. Western blots were developed using ECL2 and chemifluorescent signal was determined by a scan with the Typhoon FLA 9500 biomolecular imager. Band density of phosphorylated (4G10) and total (myc or HA) receptor was determined using ImageJ and receptor phosphorylation normalized for expression was compared between samples. Care was taken to only quantify bands within the linear range for each antibody.

## **Results**

*Transient transfection of chimeric receptors results in high background phosphorylation and inactive IGF1R-EGFR chimeric receptors*

Previous cell based IGF1R activation assays (Kavran, 2014) utilized transient transfection of HEK293 cells. Transiently transfected IGF1R in HEK293 cells has low basal phosphorylation and robustly responds to IGF1 stimulation (Figure 2.2, lanes 2 and 3). Initial attempts at chimeric receptor experiments reproduced this protocol with the intention of screening a large panel of IGF1R-RTK chimeric receptors. However, this experimental approach yielded ambiguous results. IGF1R-EGFR receptors, mimics of previously reported chimeric receptors, failed to show IGF1 induced stimulation as judged by anti-phosphotyrosine Western blots (Figure 2.2, lanes 4,5). For all chimeric receptors tested, the only chimeric IGF1R-RTK receptor displaying a robust response to ligand was the IGF1R-IR chimera (Figure 2.2, lanes 6 and 7).

Given that HEK293 cells have endogenous EGFR receptors and CHO cells lack detectable EGFR expression, transient transfection of CHO cells was tested with the same result as observed in HEK293 cells. That is, IGF1R-EGFR chimeric receptors failed to activate. This result was independent of whether the TM in the chimera was from IGF1R or EGFR or whether the ECD was IGF1R or IR. High background phosphorylation of EGFR kinase domains, both in the full length, wild-type receptor and the chimeric receptor, as a result of transient transfection was observed and cause for concern, and stable cell lines experiments were pursued to avoid potential artifacts stemming from transient transfection.

#### *Stable cell transfections reveal unique patterns of chimeric receptor activation*

Stable CHO cell lines were generated for wild-type EGFR, IR, and IGF1R, and three different chimeric receptors EGFR(ECD-TM)-IR(ICD), IR(ECD)-EGFR(TM-ICD) and IGF1R(ECD-TM)-EGFR(ICD). CHO cells lack endogenous EGFR, but stable

transfection of full length EGFR renders CHO cells responsive to EGF and robust EGFR phosphorylation is seen when EGF is added (Figure 2.3, lanes 4 to 6). Stable transfectants of full length IR and IGF1R are responsive to insulin and IGF1, respectively (Figure 2.3, lanes 2 to 4 and 14 to 16). Chimeric receptors of IR-EGFR and IGF1R-EGFR again failed to robustly respond to insulin or IGF1, respectively, as was observed in transient transfection experiments (Figure 2.3, lanes 7 to 9 and 19 to 21). A faint band may be seen for phosphorylated EGFR kinase in lanes with insulin/IGF1 addition (lanes 8 and 20); however, it is markedly less intense than stimulated wild type receptor. Whether this band is specific to the chimeric receptor but less intense will be further investigated. Given endogenous IR/IGF1R receptors that respond to ligand in CHO cells, crosstalk between ligand stimulated IR/IGF1R receptors, either directly or indirectly, could be cause of this low level phosphorylation. An IR/IGF1R small molecule kinase inhibitor will be used to test this hypothesis. In contrast to transient transfection experiments, background phosphorylation of the EGFR kinase domain in both wild type and chimeric receptors was below detection limits, and the EGFR-IR chimera displayed strong activation when stimulated with EGF (Figure 2.3, lanes 10 to 12). Given that CHO cells lack endogenous EGFR, we can rule out any signal from endogenous EGFR. The control chimeric receptor of IGF1R-IR did activate when IGF1 was added (Figure 2.3, lanes 16 to 18); however, given background levels of endogenous IGF1R, we cannot determine whether this activation is direct or the result of phosphorylation via hybrid of endogenous IGF1R heterodimerized with stably transfected IGF1R-IR receptors.

While a slight glycosylation defect is seen in the proreceptor of the IR-EGFR chimera, it is not seen in the IGF1R-EGFR chimera. Given a majority of the proreceptor

appears to be correctly processed for IR-EGFR, and the IGF1R-EGFR gives the same result, we do not believe glycosylation defects to be a cause for this negative chimera result. We will confirm that the IR-EGFR is expressed on the cell surface using cell surface biotinylation.

## **Discussion**

We were admittedly surprised by our IR-EGFR and IGF1R-EGFR chimeric receptor results and pursued multiple cell lines and stable vs. transient expression to confirm this result. For ease of discussion, we will refer to the IR-EGFR chimera but results were consistent with both IR-EGFR and IGF1R-EGFR. Initial intentions to survey activation of more classes of RTK intracellular regions when fused to IGF1R extracellular regions were set aside owing to the irreproducibility of the reported activation of IR-EGFR chimera (Riedel, 1986; Riedel, 1989). Given that our IR-EGFR chimera does signal, we are left to wonder what prevents this chimeric receptor from signaling when the inverse EGFR-IR chimera is responsive to EGF.

The EGFR and IR classes of RTKs have multiple similarities, especially when relating to extracellular domain structure and ligand binding. Both receptor families form dimers mediated solely by the receptor. For IGF1R and IR, this is the constitutive state of the receptor. For EGFR, ligand binding stimulates dimerization; however, the dimerization interface is made up solely of receptor (and not ligand) surfaces. In contrast, all other RTKs studied have fully ligand mediated dimers, or partial-ligand, partial-receptor mediated dimers. Both IR and EGFR family receptors also display negative cooperativity. ECD enforced TM association in the active state has been

proposed for both receptors based on FRET, mutagenesis and crosslinking studies, and liganded EGFR ECD crystal structures (Kavran, 2014; Lemmon, 2014)

Divergence between EGFR and IR/IGF1R is found in mechanisms of kinase domain activation. EGFR adopts an active, asymmetric dimer where the C-lobe of one kinase allosterically activates the N-lobe of the second kinase (Zhang, 2006). Activation of the IGF1R kinase domain does not involve formation of the same asymmetric dimer as in EGFR (Kavran, 2014) and active IR/IGF1R kinases have recently been suggested to adopt a symmetric, dimeric active state allowing for active loop transphosphorylation (Cabail, 2015).

Why, then, can the dimerization of the EGFR ECD activate the IR kinase domain? To evaluate possible models, we will analyze our results piecewise. First, the IR kinase domain is inactive when fused to the EGFR ECD-TM. This result is easiest to model as the EGFR ECD is a monomer, and unpublished results from Patrick Byrne of our laboratory indicated that the EGFR ECD enforces TM separation in the absence of ligand. In the absence of another kinase domain in proximity, the IR kinase domain cannot transphosphorylate.

Second, the IR kinase domain in the EGFR-IR chimera activates in response to EGF. Biochemical and structural studies of EGF bound EGFR dimers strongly supports transmembrane domains association in the active state. This proximity appears to be sufficient for the IR kinase in the EGFR-IR chimera to activate. This observation suggests that the recently suggested IGF1R mechanism of activation, where TMs apart hold the IR kinase domain inactive but bringing the TMs together or in close proximity activates the IR kinase is not dependent on IR ECD specific features (Kavran, 2014).



Conversely, why does EGFR kinase fail to activate in IR- or IGF1R-ECD chimeric fusion? First, the EGFR kinase domain is inactive when fused to the IR- or IGF1R- ECD and ligand is not present. It is assumed that the IR and IGF1R ECDs in the chimeric receptors hold the EGFR TMs apart as they do in wild-type IR/IGF1R. It is conceivable that the separation of the EGFR TMs by the IR ECD prevents the EGFR kinase from adopting the active asymmetric dimer state. Given that evidence suggests wild type EGFR exists in some fraction of preformed, inactive dimers, it is probable that an inactive, dimeric state of the EGFR kinase exists, and this inactive state is favored in chimeric receptors of the IR or IGF1R ECD and the EGFR TM-ICD.

Second, addition of insulin to the IR-EGFR chimera or IGF1 to the IGF1R-EGFR chimera fails to strongly activate the intracellular EGFR kinase. This result is the most perplexing and requires experimental follow-up and further thought. Ligand binding to the IR/IGF1R ECD results in a conformational change that brings the TMs in association. Whether this same change occurs in the chimeric receptor has not yet been determined. On one hand, if the conformational change is the same and TMs are brought together, this would suggest that the TM positioning in a liganded IR/IGF1R receptor is incompatible with formation of the active, asymmetric EGFR kinase domain dimer. Given the reliance of EGFR activation on this asymmetric dimer, the inability to adopt this dimer would prevent receptor phosphorylation and rationalize our results. To test whether the IR/IGF1R ECD conformational change occurs in the chimeric receptor, a mutation in the extracellular JM region, H905C, which traps the TMs together and activates the IR/IGF1R kinase domains, will be made and tested for activity.

Another possibility is that an interaction of the inactive kinases in the absence of ligand is energetically favorable, and the energy required to transition from inactive to active state is greater than that generated by the IGF1R or IR ECD upon ligand binding. Mutations in the EGFR kinase domain, such as L858R, that lead to ligand independent kinase activation, may aid in parsing these models when introduced into the chimeric receptor.

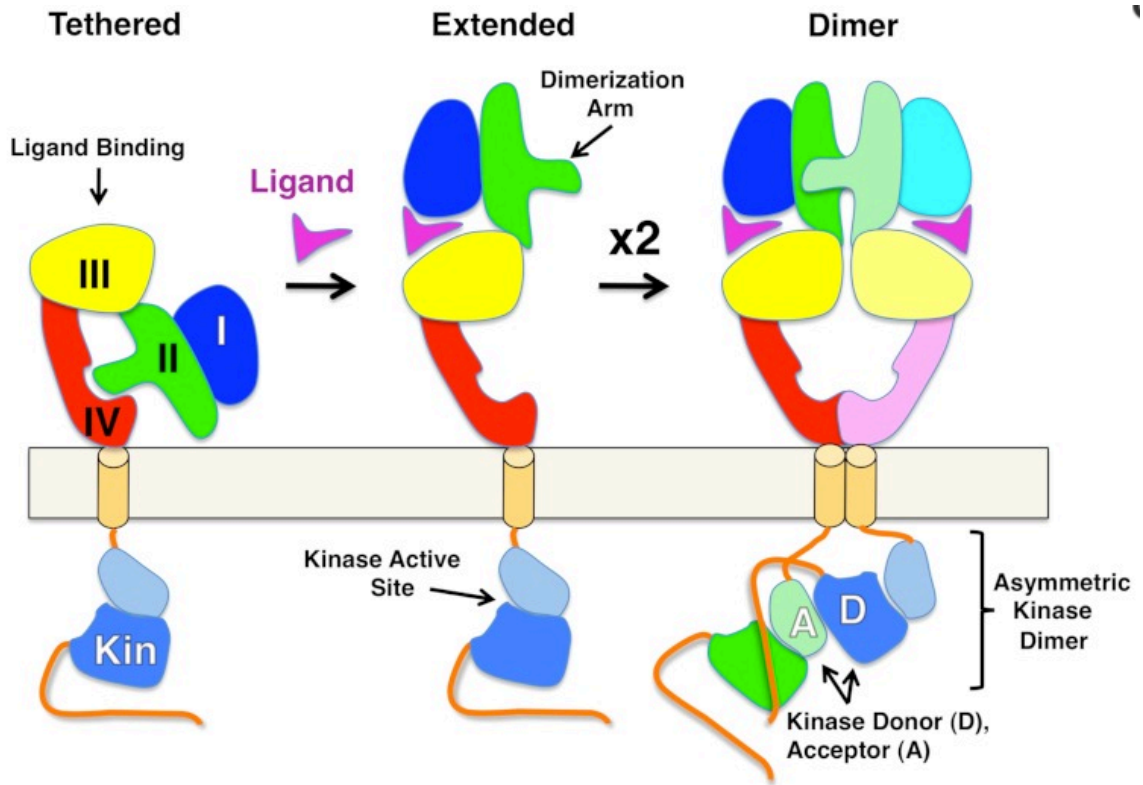
Because we were unable to directly replicate previous chimeric receptor work, we question whether the relatively small (two- to four- fold) changes in autoradiography images with very high background may have been over interpreted (Riedel, 1989). Additionally, the antibodies they used to immunoprecipitate the receptors may have selectively immunoprecipitated ligand bound receptor. Given the nature of the experiment and data presented in the earlier work, it is not possible to discern the total amount of protein in each post-immunoprecipitation kinase assay. We have made every effort to reproduce the chimeric receptors used in previous reports and to use a cell line, CHO cells, which lacks at least one of the receptors of interest, EGFR.

Future experiments include analysis of cell lysates to determine the activation state of downstream effector proteins. Given that both the IGF1R/IR and EGFR pathways feed into the PI3K/MAPK pathway, it may be difficult to distinguish activation of the pathway by ligand stimulation of endogenous receptors vs. stably transfected clones. We hope that a low background will allow us to confirm that insulin/IGF1 addition to the IR- or IGF1R-EGFR chimeric receptors fails to activate downstream pathway components when compared with the EGFR full length stable cell line.

Mutational analysis of the IGF1R-EGFR chimeric receptor may also prove useful. As described above, the H905C mutation may answer whether the TMs do associate in the IR-EGFR and IGF1R-EGFR chimeras and if this TM association activates the EGFR kinase. Additional information may come from decoupling the IGF1R ECD from the EGFR TM-ICD by insertion of flexible linkers of variable lengths, which activated the wild type IGF1R with as little as 5-10 amino acids inserted (Kavran, 2014). Work with EGFR has also shown that adding residues to the extracellular juxtamembrane region decouples kinase activation from ECD inhibition (Chung, 2010). If the specific TM conformation by the ligand bound IR/IGF1R-EGFR chimeric receptors inhibits the EGFR active state conformation, loosening the linkage may allow the TMs to assume a conformation compatible with the active, asymmetric dimer.

What then to do with question of whether the IGF1R(ECD-TM)-RTK(ICD) chimeric receptors made with other RTK classes are responsive to IGF1? Results in transient transfections of HEK293 cells indicated that no other chimeras were active. First, stable cell lines, which provide the most reliable results, must be generated for both the chimeric and full length receptors of the RTKs of interest. Creation of a small panel of these covering well known RTK classes such as Class III and Class IV where ligand binding is thought to bring membrane proximal extracellular domains in proximity is underway. Chimeric receptors with ectodomains of IGF1R and EGFR should be assayed side by side. While the IGF1R ectodomain may prove more specialized, given the EGFR-IR positive result, the monomer-dimer transition upon ligand binding to EGFR may be more likely to activate non-cognate RTK kinase domains.

Our data suggest that the manner in which IGF1 and insulin activate their receptors does not activate the EGFR kinase domain in chimeric receptors. However, the manner in which EGFR activates its kinase domain does apparently activate the IR kinase. This dichotomy between ECD identity and kinase activation suggests perhaps the insulin receptor family of RTKs, unique dimers amongst all RTK classes, have a specialized mechanism of activation that fails to cross to other RTK classes.



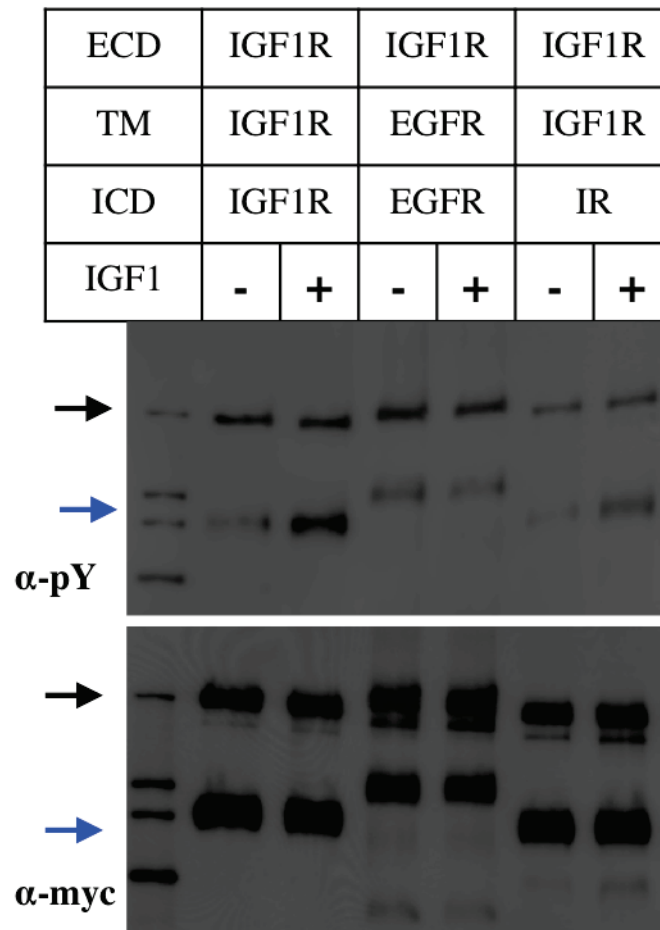
**Figure 2.1 Epidermal growth factor activation.** In the absence of ligand, EGFR adopts a tethered conformation. Ligand binding to the extracellular domain of EGFR results in the adoption of an extended conformation and subsequent dimerization. The kinase domains of dimeric EGFR adopt a specific conformation mediated by an asymmetric kinase dimer. Figure reproduced from (Liu, 2012).

**Table 2.1 RTK Expression Vectors.** A set of mammalian expression vectors were created from parent vector pαH. Vector multiple cloning sites (MCS) were adapted to remove the signal sequence (V1), add C-terminal tandem myc-HA epitope tags (all), or add various RTK ECD constructs (V3-V7) for fusion to alternative RTK ICDs

| RTK Vector Multiple Cloning Sites |   |
|-----------------------------------|---|
| Vector                            | Description   |
| pαH                               | <p>M G I L P S P G M P A L L S L V S L</p> <p><u>GGTACC</u> <u>GCCACC</u> <b>ATGGGCATCCTTCCCGAGCCCTGGGATGCCTGCGCTGCTCTCCCTCGTGAGCCTT</b></p> <p>KpnI Kozak PTPα signal sequence</p> <p>L S V L L M G C V A E G S S R T R S G R</p> <p><u>CTCTCCGTGCTGCTGATGGGTTCGCTAGCTGAA</u> <u>GGATCC</u> <u>TCTAGA</u> <u>ACGCGT</u> <u>AGCGGCCGC</u></p> <p>BamHI XbaI MluI NotI</p> <p>H E F H H H H H H H H * L E</p> <p>CAT <u>GAATTC</u> CATCACCACCATCATCACCATCAT TGA <u>CTCGAG</u></p> <p>EcoRI 8xHisTag XhoI</p>   |
| V1                                | <p>G T G S S R T R S G R H E F E Q K L I S</p> <p><u>GGTACC</u> <u>GGATCC</u> <u>TCTAGA</u> <u>ACGCGT</u> <u>AGCGGCCGC</u> CAT <u>GAATTC</u> <u>GAGCAGAAGCTGATCAG</u></p> <p>KpnI BamHI XbaI MluI NotI EcoRI myc</p> <p>E E D L Y P Y D V P D Y A * L E</p> <p><u>CGAGGAGGACCTG</u> <u>TACCCTTACGACGTTCCCTGACTACGCC</u> TGA <u>CTCGAG</u></p> <p>HA XhoI</p>  |
| V2                                | <p>M G I L P S P G M P A L L S L V S L</p> <p><u>GGTACC</u> <u>GCCACC</u> <b>ATGGGCATCCTTCCCGAGCCCTGGGATGCCTGCGCTGCTCTCCCTCGTGAGCCTT</b></p> <p>KpnI Kozak PTPα signal sequence</p> <p>L S V L L M G C V A E G S S R T R S G R</p> <p><u>CTCTCCGTGCTGCTGATGGGTTCGCTAGCTGAA</u> <u>GGATCC</u> <u>TCTAGA</u> <u>ACGCGT</u> <u>AGCGGCCGC</u></p> <p>BamHI XbaI MluI NotI</p> <p>H E F E Q K L I S E E D L Y P Y D V P D Y</p> <p>CAT <u>GAATTC</u> <u>GAGCAGAAGCTGATCAGCGAGGAGGACCTG</u> <u>TACCCTTACGACGTTCCCTGACTAC</u></p> <p>EcoRI myc HA</p> <p>A * L E</p> <p><u>GCC</u> TGA <u>CTCGAG</u></p> <p>XhoI</p> |
| V3                                | <p>as in MCS_B with the following insertion noted in <b>bold</b>:</p> <p>M K S I H L S R</p> <p><u>GGTACC</u> <u>GGATCC</u> <b><u>GCCACC ATGAAGTCT // ATCCATCTG TCTAGA</u></b></p> <p>KpnI BamHI Kozak IGF1R(ECD, 1-936) XbaI</p>   |
| V4                                | <p>as in MCS_B with the following insertion noted in <b>bold</b>:</p> <p>M K S V F H S R</p> <p><u>GGTACC</u> <u>GGATCC</u> <b><u>GCCACC ATGAAGTCT // GTCTTCCAT TCTAGA</u></b></p> <p>KpnI BamHI Kozak IGF1R(ECD-TM, 1-959) XbaI</p>  |
| V5                                | <p>as in MCS_B with the following insertion noted in <b>bold</b>:</p> <p>M K S A A D S R</p> <p><u>GGTACC</u> <u>GGATCC</u> <b><u>GCCACC ATGAAGTCT // GCTGCTGAT TCTAGA</u></b></p> <p>KpnI BamHI Kozak IGF1R(ECD-iJM, 1-985) XbaI</p>   |
| V6                                | <p>as in MCS_C with the following insertion noted in <b>bold</b>:</p> <p>V A E G S S R H L Y P S N S R T R</p> <p><u>GTAGCTGAA</u> <u>GGATCC</u> <u>TCTAGA</u> <b><u>CACCTGTAC // CCGTCAAAT TCTAGA ACGCGT</u></b></p> <p>PTPαSS BamHI XbaI IR(ECD, 28-953) XbaI MluI</p>  |
| V7                                | <p>as in MCS_B with the following insertion noted in <b>bold</b>:</p> <p>M R P L F M S R</p> <p><u>GGTACC</u> <b><u>GCCACC ATGCGACCC // CTCTTCATG TCTAGA</u></b></p> <p>KpnI Kozak EGFR(ECD-TM, 1-668) XbaI</p>   |

**Table 2.2 Panel of RTK Clones.** For each RTK, a subset of clones was made to express full length, truncated and chimeric receptors with C-terminal myc-HA epitope tags.

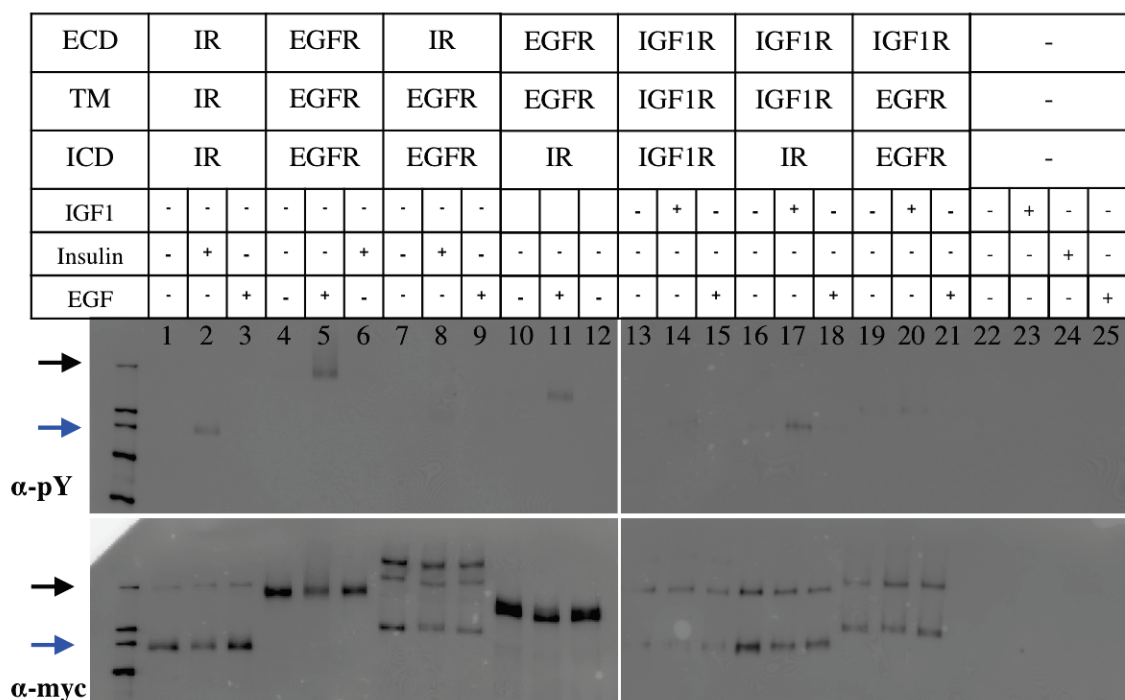
| #  | RTK    | Name | Expresses:                      | #  | RTK   | Name | Expresses:               |
|----|--------|------|---------------------------------|----|-------|------|--------------------------|
| 1  | EGFR   | 1A   | EGFR                            | 13 | FGFR2 |      | FGFR2                    |
|    |        | 1B   | EGFR(TM-ICD)                    |    |       | 13B  | FGFR2(TM-ICD)            |
|    |        | 1C   | Cys-EGFR(TM-ICD)                |    |       | 13C  | Cys-FGFR2(TM-ICD)        |
|    |        | 1E   | IGF1R(ECD)-EGFR(TM-ICD)         |    |       | 13E  | IGF1R(ECD)-FGFR2(TM-ICD) |
|    |        | 1F   | IGF1R(ECD-TM)-EGFR(ICD)         | 15 | Ptk7  | 15A  | Ptk7                     |
|    |        | 1G   | IGF1R(ECD-TM-iJM)-EGFR(KD-tail) |    |       | 15B  | Ptk7(TM-ICD)             |
|    |        | 1H   | IR(ECD)-EGFR(TM-ICD)            |    |       | 15C  | Cys-Ptk7(TM-ICD)         |
|    |        |      |                                 |    |       | 15E  | IGF1R(ECD)-Ptk7(TM-ICD)  |
| 2  | ErbB2  | 2A   | ErbB2                           | 16 | TrkA  | 16A  | TrkA                     |
|    |        | 2B   | ErbB2(TM-ICD)                   |    |       | 16B  | TrkA(TM-ICD)             |
|    |        | 2C   | Cys-ErbB2(TM-ICD)               |    |       | 16C  | Cys-TrkA(TM-ICD)         |
|    |        | 2E   | IGF1R(ECD)-ErbB2(TM-ICD)        |    |       | 16E  | IGF1R(ECD)-TrkA(TM-ICD)  |
| 3  | ErbB3  | 3A   | ErbB3                           |    |       | 16F  | IGF1R(ECD-TM)-TrkA(ICD)  |
|    |        | 3B   | ErbB3(TM-ICD)                   | 19 | MuSK  | 19A  | MuSK                     |
|    |        | 3C   | Cys-ErbB3(TM-ICD)               |    |       | 19B  | MuSK(TM-ICD)             |
|    |        | 3E   | IGF1R(ECD)-ErbB3(TM-ICD)        |    |       | 19C  | Cys-MuSK(TM-ICD)         |
| 5  | IR     | 5A   | IR                              |    |       | 19E  | IGF1R(ECD)-MuSK(TM-ICD)  |
|    |        | 5B   | IR(TM-ICD)                      |    |       | 19F  | IGF1R(ECD-TM)-MuSK(ICD)  |
|    |        | 5C   | Cys-IR(TM-ICD)                  | 20 | Met   | 20A  | Met                      |
|    |        | 5E   | IGF1R(ECD)-IR(TM-ICD)           |    |       | 20B  | Met(TM-ICD)              |
|    |        | 5F   | IGF1R(ECD-TM)-IR(ICD)           |    |       | 20C  | Cys-Met(TM-ICD)          |
|    |        | 5G   | IGF1R(ECD-TM-iJM)-IR(KD-tail)   |    |       | 20E  | IGF1R(ECD)-Met(TM-ICD)   |
|    |        | 5H   | EGFR(ECD-TM)-IR(TM-ICD)         |    |       | 20F  | IGF1R(ECD-TM)-Met(ICD)   |
| 6  | IGF1R  | 6A   | IGF1R                           | 21 | Ron   | 21A  | Ron                      |
|    |        | 6B   | IGF1R(TM-ICD)                   |    |       | 21B  | Ron(TM-ICD)              |
|    |        | 6C   | Cys-IGF1R(TM-ICD)               |    |       | 21C  | Cys-Ron(TM-ICD)          |
|    |        |      |                                 |    |       | 21E  | IGF1R(ECD)-Ron(TM-ICD)   |
| 7  | PDGFR  | 7A   | PDGFR                           |    |       | 21F  | IGF1R(ECD-TM)-Ron(ICD)   |
|    |        | 7B   | PDGFR(TM-ICD)                   | 23 | Tyro3 | 23A  | Tyro3                    |
|    |        | 7C   | Cys-PDGFR(TM-ICD)               |    |       | 23B  | Tyro3(TM-ICD)            |
|    |        | 7E   | IGF1R(ECD)-PDGFR(TM-ICD)        |    |       | 23C  | Cys-Tyro3(TM-ICD)        |
| 10 | Flt1   | 10A  | Flt1                            |    |       | 23E  | IGF1R(ECD)-Tyro3(TM-ICD) |
|    |        | 10B  | Flt1(TM-ICD)                    | 27 | EphA3 | 27A  | EphA3                    |
|    |        | 10C  | Cys-Flt1(TM-ICD)                |    |       | 27B  | EphA3(TM-ICD)            |
|    |        | 10E  | IGF1R(ECD)-Flt1(TM-ICD)         |    |       | 27C  | Cys-EphA3(TM-ICD)        |
|    |        |      |                                 |    |       | 27E  | IGF1R(ECD)-EphA3(TM-ICD) |
| 11 | VEGFR2 | 11A  | VEGFR2                          |    |       |      |                          |
|    |        | 11B  | VEGFR2(TM-ICD)                  |    |       |      |                          |
|    |        | 11C  | Cys-VEGFR2(TM-ICD)              |    |       |      |                          |
|    |        | 11E  | IGF1R(ECD)-VEGFR2(TM-ICD)       |    |       |      |                          |
|    |        |      |                                 |    |       |      |                          |
| 12 | FGFR1  | 12A  | FGFR1                           |    |       |      |                          |
|    |        | 12B  | FGFR1(TM-ICD)                   |    |       |      |                          |
|    |        | 12C  | Cys-FGFR1(TM-ICD)               |    |       |      |                          |
|    |        | 12E  | IGF1R(ECD)-FGFR1(TM-ICD)        |    |       |      |                          |
|    |        |      |                                 |    |       |      |                          |



**Figure 2.2 Transiently transfected IGF1R-EGFR chimeric receptors fail to activate.**

HEK293 cells transiently transfected with C-terminal myc-HA tagged IGF1Rfl (lanes 2, 3), IGF1R(ECD)-EGFR(TM-ICD) (lanes 4, 5) and IGF1R(ECD-TM)-IR(ICD) were serum starved and then stimulated with 20nM IGF1 where indicated. Normalized lysates were immunoprecipitated with an anti-HA antibody and immunoblotted with anti-phosphotyrosine (4G10) and anti-myc (9E10) antibodies. The black arrow denotes approximate position of proreceptor and the blue arrow denotes approximate position of the Beta subunit.





**Figure 2.3 IGF1R and EGFR activation in CHO cell stable cell lines.** CHO stable cells for each construct indicated (from left to right, IR-full length, EGFR-full length, IR(ECD)-EGFR(TM-ICD), EGFR(ECD-TM)-IR(ICD), IGF1R-full length, IGF1R(ECD-TM)-IR(ICD), IGF1R(ECD)-EGFR(TM-ICD), all containing C-terminal myc-HA epitope tags) were serum starved for three hours and then stimulated with indicated ligand for 5 minutes. Normalized cell lysates were immunoprecipitated and immunoblotted with anti-phosphotyrosine (4G10) and anti-myc (9E10) antibodies. The black arrows denote approximate position of proreceptor and the blue arrows denote approximate position of the Beta subunit.

### **Chapter 3. Smoothened Structure and Function**

#### **Acknowledgments**

The introduction of this chapter is published as a minireview in the *Journal of Biological Chemistry* by Jacqueline McCabe and Daniel J. Leahy (McCabe, 2015). Work found in the methods, results and discussion section was performed in close collaboration with Ben Myers and Phil Beachy of Stanford University.

#### **Abstract**

A general aim of studies of signal transduction is to identify mediators of specific signals, order them into pathways, and understand the nature of interactions between individual components and how these interactions alter pathway behavior. Despite years of intensive study and its central importance to animal development and human health, our understanding of the Hedgehog (Hh) signaling pathway remains riddled with gaps, question marks, assumptions, and poorly understood connections. In particular, understanding how interactions between Hh and Patched (Ptc), a 12-pass integral membrane protein, lead to modulation of the function of Smoothened (Smo), a 7-pass integral membrane protein, has defied standard biochemical characterization. Recent structural and biochemical characterizations of Smoothened domains have begun to unlock this riddle, however, and point towards improved cancer therapies. Building on this evidence, we set out to express, purify, biochemically characterize and determine the structure of a constructs containing N-terminal cysteine rich domain (CRD) and 7-pass

transmembrane portion (7TM) of Smo utilizing a panel of CRD and 7TM small molecule modulators.

## **Introduction**

Members of the Hedgehog (Hh) family of secreted signaling proteins are present in most metazoans and owe their name to the effects loss of Hh function has on *Drosophila* embryos, which lose their normal segmented pattern and develop a uniform coat of bristles reminiscent of the coats of hedgehogs (Nusslein-Volhard, 1980). As presaged by this phenotype, Hh proteins mediate essential tissue patterning events during many stages of animal development (Ingham, 2011), and abnormal Hh function is associated with birth defects and cancer (Briscoe, 2013). Hh proteins are also involved in tissue maintenance and wound repair in adult animals (Arwet, 2012). Hh proteins achieve their patterning effects by functioning as classical morphogens (Heemsker, 1994). That is, Hh proteins form gradients of decreasing concentration from sites of secretion and induce concentration-dependent differentiation of distinct cell types (Gradilla, 2013; Ribes, 2009). As befits a morphogen, Hh expression, release, diffusion, and signal reception are tightly regulated by multiple factors (Therond, 2012).

Classical and modern genetic techniques have identified several cell-surface proteins and glycans involved in receiving or modifying Hh signals (Ryan, 2012). The core components of this process, conserved in all organisms known to have active Hh signaling, are Patched (Ptc) and Smoothed (Smo) (Figure 3.1) (Nakano, 1989; Stone, 1996; Alcedo, 1996; van den Heuvel, 1996). Ptc functions upstream of Smo and has been genetically and biochemically defined as a primary component of the Hh receptor

(Zheng, 2010; Fuse, 1999). Ptc is a 12-pass integral membrane protein with distant homology to bacterial Resistance-Nodulation-Cell Division (RND) transporters (Taipale, 2002; Tsent, 1999). Transmembrane helices 2-6 of Ptc are also homologous to sterol-sensing domains (SSDs), which are found in diverse integral membrane proteins and regulate activity in response to levels of free cellular sterols (Kuwabara, 2002). Smo is a member of the Frizzled family (class F) of G-protein coupled receptors (GPCRs) (Ayer, 2010), and contains an N-terminal, ~14kDa extracellular Cysteine-Rich Domain (CRD) connected via a linker to 7 membrane-spanning helices (7TM) and an extended (~200 amino acids, human; ~450 amino acids, *Drosophila*) C-terminal tail.

Hh signaling responses are modulated by additional cell-surface components including CDO, BOC, Gas1, Hedgehog-interacting Protein, and Glypicans in vertebrates and Ihog, BOI, and the glypican Dally-like Protein in flies (Beachy, 2010; Yao, 2006; McLellan, 2006; McLellan, 2008, Bosanac, 2009; Bishop, 2009; Williams, 2010; Kim, 2011; Izzi, 2011; Allen, 2011). These factors either lack intracellular regions owing to glycosylphosphatidylinositol anchors (Gas1, Glypicans) or have intracellular regions that are not implicated in Hh signaling and do not appear to transmit Hh signals across the cell membrane directly (Zheng, 2010). Instead, transmission of Hh signals across the membrane appears to be mediated by Smo, the most downstream cell-surface component of the Hh signaling pathway. Consistent with this role, the cytoplasmic tail of Smo becomes heavily phosphorylated and likely changes disposition when the Hh pathway is active (Zhang, 2004; Jia, 2004; Zhao, 2007). These changes are coupled to intracellular signaling events that ultimately converge on members of the Gli family of transcription

factors, active forms of which translocate to the nucleus and upregulate expression of target genes (Hui, 2011).

Recent discovery of the importance of Ptc and Smo localization for normal Hh signaling has added additional complexity to Hh pathway regulation. In vertebrates, Sonic Hh (Shh) and Hh pathway agonists result in movement of Smo from the plasma membrane to the primary cilium, a nonmotile flagellar-like organelle present on most cells, and dispersal of Ptc from its previous localization at the base of the primary cilium (Rohatgi, 2007). Although movement of Smo to the primary cilium appears essential for normal Hh signaling in vertebrates (Haungfu, 2003), this movement is neither sufficient for signaling (Rohatgi, 2009) nor conserved in flies (Denef, 2000), and a core signaling capacity that is independent of ciliary localization must be present in Smo. This minireview will focus on recent advances in structural and biochemical characterization of Smo, and readers are encouraged to consult other sources for background on additional Hh pathway components.

#### *Patched and Smoothened*

In unstimulated Hh-responsive cells Ptc functions upstream of Smo to inhibit its activity (Ingham, 2011). Hh triggers signaling responses by interacting with Ptc to relieve this inhibition, but both how Ptc inhibits Smo and how Hh relieves this inhibition remain unclear. As a small amount of Ptc is sufficient to inhibit a large stoichiometric excess of Smo (Taipale, 2002), Ptc does not appear to inhibit Smo through a direct interaction. Rather, the homology of Ptc to transporters and the ability of Smo activity to be modulated by small molecules have led to the widespread belief that Ptc controls Smo through transport of a small molecule intermediary (Taipale, 2002). Indeed, the ability of

Smo to bind and be inhibited by the plant sterol cyclopamine led to the development of compounds targeting the cyclopamine binding site for the treatment of cancers with abnormally active Hh signaling (Cooper, 1998; Taipale, 2000). As some Smo-binding compounds function as Hh pathway agonists, it has been tempting to speculate that an endogenous cyclopamine-like compound modulates Smo activity (Chen, 2002). Indeed, the sterol vitamin D3 has been proposed to function as a Ptc-dependent inhibitor of Smo (Bijlsma, 2006), although this observation awaits confirmation.

#### *Smo: 7TM Region*

The absence of knowledge of the physiological factors responsible for Smo activation—or inhibition—has presented a frustrating barrier to understanding Hh pathway regulation, but several recent results have begun to clarify this issue. Firstly, Stevens and colleagues have determined atomic-resolution crystal structures of the 7TM region of human Smo complexed with 5 different small molecules, including cyclopamine (Wang, 2013; Wang, 2014; Weierstall, 2014). These landmark structures show that, despite sharing less than 10% sequence identity with class A GPCRs such as rhodopsin and the  $\beta_2$ -adrenergic receptor ( $\beta_2$ AR), the Smo TM regions adopt an overall conformation very similar to that of inactive class A GPCRs (Figure 3.2A) (Venkatakrisnan, 2013). As discussed in more detail below (Xie, 1998), this structural homology couples with the observation that activating mutation in Smo occur at sites that appear to stabilize the inactive state of class A GPCRs to suggest that the 7TM region of Smo is likely to undergo GPCR-like conformational changes during its activity cycle (Venkatakrisnan, 2013). Such a conformational cycle would also be consistent with the

ability of Smo to signal through G-proteins in certain circumstances (Riobo, 2006; Philipp, 2009; Polizio, 2011a; Polizio, 2011b; Shen, 2013; Carbe, 2014).

Although the overall fold of its 7TM bundle is conserved with other GPCRs, Smo has additional features including an extension of extracellular loops (ECLs). All of the co-crystallized compounds bind Smo in a long narrow pocket formed by the ECL extensions and upper portions of the transmembrane domains (Figure 3.2) (Wang, 2013; Wang, 2014; Weierstall, 2014). The drug-binding pocket is exposed to the extracellular space suggesting that drugs and any endogenous ligands access the pocket from the extracellular surface. This extracellular accessibility contrasts with a class A GPCR lipid receptor where the extracellular loops form a closed cage and ligand is thought to access its binding site from within the membrane (Hanson, 2012). Although the CRD was deleted from the crystallized Smo 7TM domain, the majority of the residues of the extracellular linker between the CRD and the first TM domain are present and adopt an ordered structure. Disulfide bonds both within the linker and between the linker and the second extracellular loop appear to stabilize the linker structure (Figure 3.2A), and disruption of these disulfides results in increased Smoothed activity (Carroll, 2012). In addition, the extracellular linker interacts with the extended extracellular loop connecting TMs VI and VII (ECL3), which forms a cap over the drug-binding pocket. This ordered linker region suggests that the CRD may be directly coupled to the 7TM region and influence its conformation.

The five compounds crystallized in the Smo binding pocket include an agonist (SAG1.5) and four antagonists (LY2940680, SANT1, ANTA XV, and cyclopamine) (Figure 3.4). All ligands bind in the pocket with their long axes perpendicular to the

plane of the membrane but vary in their depth relative to the extracellular outlet (Figure 3.2B). At the extremes, cyclopamine interacts predominantly with the extracellular loops while another antagonist, SANT-1, binds deep within the pocket, which spans 28 Å from the top of cyclopamine to the bottom of SANT1. Asp 473, a residue that when mutated to histidine confers resistance to the anti-cancer agent Vismodegib (GDC-0449) (Yauch, 2009; Robarge, 2009), lines the drug-binding pocket but interacts differently with different antagonists and does not confer universal drug resistance (Wang, 2013). Asp 473 does not directly contact LY2940680, for example, and the D473H substitution does not affect the activity of LY2940680 (Bender, 2011). The variable susceptibility of individual drugs to resistance mutations suggests that second generation drugs or combination therapies may prolong the time to development of resistance.

LY2940680, cyclopamine, Anta XV, and the agonist SAG1.5 contact the Smoothed extracellular loops lining the top of the ligand binding cavity, but SANT1 binds more deeply in the pocket and only contacts ECL2, which is positioned within the 7TM region. In contrast to cyclopamine, which binds more tightly to Smo than to a constitutively active Smo variant bearing a single-site substitution (SmoM2), SANT1 binds both Smo and SmoM2 with equal potency (Chen, 2002). How the position of SANT1 deep within the 7TM bundle correlates to its ability to bind both Smo and SmoM2, whose W535L substitution occurs at the base of TM VII, is not clear. Also of interest are the variable effects Smo antagonists have on Smo localization. SANT1, LY2940680, and cyclopamine all inhibit Smo function, but only cyclopamine promotes the translocation of a still-inactive Smo to the primary cilium indicating translocation and activation are separable functions.



The failure of Smo to adopt an active-like conformation when bound to the agonist SAG1.5 is curious but not unprecedented for agonist-bound GPCRs (Warne, 2011). Binding of an agonist to an apparently inactive state may reflect a low energetic barrier between active and inactive states, conformational flexibility of the active state (Nygaard, 2013), and/or the effects of truncation of Smo N- and C-terminal regions. SAG1.5 binds in the same region of the binding pocket as LY2940680, ANTA XV, and cyclopamine, and Smo with SAG1.5 bound displays only slight alterations in binding pocket residues. Larger conformational changes associated with active state GPCRs, such as the movements of TMs VI and VII to accommodate G-protein binding, are not seen in the Smo-SAG1.5 structure. Crystallization of an active state of Smo may require adding back the CRD or portions of the C-terminal tail or co-crystallization with active conformation specific nanobodies (Ring, 2013). Interesting subtleties about the effects of these different drugs on the conformational equilibria of intact Smo and their relation to Smo function clearly remain to be worked out.

#### *Smoothened: Cysteine-Rich Domain*

A second major insight into Smo regulation emerged when three groups independently showed that oxysterols, oxidized derivatives of cholesterol, bind specifically to the Smo CRD and activate the Hh signaling pathway (Myers, 2013; Nedelcu, 2013; Nachtergaele, 2013). Oxysterol binding by the Smo CRD is functionally as well as physically separable from small molecule binding to the 7TM site as deletion of the Smo CRD results in loss of oxysterol activation of Smo but does not affect the function of agonists and antagonists that target the 7TM region (Myers, 2013). It had previously been shown that oxysterols could modulate Hh signaling by affecting Smo

function (Corcoran 2006; Dwyer, 2007; Nachtergaele, 2012). The site of oxysterol action was not characterized at that time, although oxysterols did not appear to compete with cyclopamine for binding to Smo (Nachtergaele, 2012).

The new reports all show that 20(S)-hydroxycholesterol (20(S)-OHC) (Figure 3.4B) activates Smo by binding to the CRD. Additionally, the Rohatgi and Siebold groups were able to determine the crystal structure of the zebrafish Smo CRD (Figure 3.3B) (Nachtergaele, 2013). All groups mapped the site of sterol action on the CRD via mutagenesis and *in silico* modeling to a hydrophobic groove that is homologous to the site at which the palmitoyl group of Wnt binds to the Frizzled CRD (Figure 3.3A) (Myers, 2013; Nachtergaele, 2013; Janda, 2012), confirming an earlier prediction based on structural homology that this region of Smo and Frizzled CRDs would bind lipophilic molecules (Bazan, 2009). Curiously, the *Drosophila* Smo CRD does not bind to 20(S)-OHC (Myers, 2013), but it and human Smo CRD were recently shown to bind to the glucocorticoid budesonide (Figure 3.4B) suggesting that sterol binding by the Smo CRD may be a conserved feature of Hh signaling (Rana, 2013). Glucocorticoids represent an interesting class of Smo modulators as both inhibitors and activators of the Hh pathway have been found with glucocorticoid scaffolds, and budesonide inhibits wt Smo, SmoD473H and SmoM2 equally well - ideal features for a Smo-targeting drug (Wang, 2012).

Variable specificity for 20(S)-OHC among Smo CRDs is perhaps not surprising given that the absence of a cellular sterol hydroxylase known to produce it makes it unlikely to be an endogenous ligand (Meyers, 2013). Assuming endogenous ligands for Smo CRDs exist, the question naturally arises of what that ligand is. A survey of

oxysterols for Smo modulatory activity found that 7-keto-27-OHC and 7-keto-25-OHC, both metabolites of 7-ketocholesterol, are able to stimulate Hh signaling in a manner that depends on the presence of the CRD (Myers, 2013). Compounds that bind the CRD and inhibit (azasterols, e.g 22-NHC) (Figure 3.4B) or partially agonize (20(R)-yne, 20-keto-yne) Smo activity validates the Smo CRD as a potential drug target and raises the possibility that an endogenous ligand for the Smo CRD may be an inhibitor rather than an activator (Nedelcu, 2013; Nachtergaele, 2013). More work is needed to identify and validate potential CRD ligands, but it seems likely that such ligands exist, and their discovery and characterization will take our understanding of Hh pathway regulation in new directions.

Two immediately exciting prospects stimulated by the discovery of a specific and functionally important sterol-binding site on the Smo CRD were that it might be route by which Ptc modulates Smo activity or rationalize why cholesterol depletion reduces Hh signaling (Cooper, 2003). Defying Occam's razor, however, oxysterol binding by the Smo CRD cannot fully account for either of these processes. Deletion of the CRD from Smo ( $\Delta$ CRDSmo) alters but does not abolish Shh mediated pathway activation (Myers, 2013; Nedelcu, 2013; Nachtergaele, 2013). Although varying levels of responsiveness of  $\Delta$ CRDSmo to Shh have been reported, this is likely due to varying tags and expression systems. Rohatgi and colleagues showed that oxysterol binding mutants of Smo retain negative regulation by Ptc and respond to Shh (Nachtergaele, 2013). Beachy and colleagues showed that  $\Delta$ CRDSmo has higher basal activity, but this activity can be reduced by co-expression with Ptc and further enhanced by addition of Shh indicating that Ptc can exert its effects on Smo independent of the CRD (Myers, 2013). Higher

basal activity and Shh responsiveness of  $\Delta$ CRDSmo were also reduced by cyclodextrin depletion, which reduces wild-type Smo activity (Cooper, 2003), suggesting the CRD is also not essential for this process but rather that cholesterol within the cell membrane is needed for normal Smo function. Indeed, a specific role for membrane-localized cholesterol in Smo modulation has been suggested (Myers, 2013) although no ordered cholesterol molecules were identified in the Smo crystal structures. Modulation of Smo activity independent of the CRD or cyclopamine-binding pocket is not unprecedented as Itraconazole (Figure 3.4C) acts on Smo at a site distinct from both the canonical 7TM pocket and the CRD to inhibit Hh pathway activity (Kim, 2013).

Cholesterol binding to the 7TM region of GPCRs is also not unprecedented. The structure of  $\beta_2$ AR bound to cholesterol and the partial inverse agonist timolol led Stevens and colleagues to propose a cholesterol consensus motif (CCM) in class A GPCRs (Hanson, 2008). The CCM comprises 3 residues predictive of cholesterol binding: an aromatic residue (W,Y) at position 4.50, a positively charged residue at or about position 4.43 that interacts with the cholesterol hydroxyl group, and a hydrophobic residue at position 4.46. The positions here refer to the Ballesteros-Weinstein numbering for GPCRs (Ballesteros, 1995), which allows cross comparison of topologically equivalent residues in GPCR TMs and was recently extended to class F GPCRs (Wang, 2013). Interestingly, Smo W365<sup>4.50</sup> overlays well with  $\beta_2$ AR W158<sup>4.50</sup>, which stacks against the sterol ring of cholesterol in the  $\beta_2$ AR structure with cholesterol bound. While Smo does not have a positively charged residue at position 4.43, Smo residue H361<sup>4.46</sup> maps to the hydrophobic position 4.46 of the CCM. A nitrogen on the imidazole ring of H361<sup>4.46</sup> is within 3.6 Å of the cholesterol hydroxyl group from cholesterol bound  $\beta_2$ AR structure

(Hanson, 2008). Whether these highly conserved class F residues, W365<sup>4.50</sup> and H361<sup>4.46</sup>, act as an alternative cholesterol-binding motif presents an intriguing possibility.

### *Targeting Smoothed in the Clinic*

Hh pathway activating mutations in the gene encoding Ptc, and less commonly the gene encoding Smo, are found in subsets of several cancers, most notably basal cell carcinoma (BCC) and pediatric medulloblastomas (Xie, 1998; Goodrich, 1997). Constitutively active mutants (CAMs) of Smo found in sporadic BCC (W535L<sup>7.55</sup> “SmoM2”) and more recently in meningiomas and ameloblastomas (W535L<sup>7.55</sup>, L412F<sup>5.51</sup>) are resistant to Vismodegib treatment (Xie, 1998; Sweeny, 2014; Brastianos, 2013; Clark, 2013). Superscripts refer to B&W numbering. W535<sup>7.55</sup> is absolutely conserved in class F GPCRs and maps to the intracellular tip of TM VII, a region structurally homologous to the NPxxY motif in class A GPCRs (Barak, 1995; Rosenbaum, 2009). W535<sup>7.55</sup> overlays with the Y<sup>7.53</sup> of the NPxxY motif, which undergoes rearrangement in inactive vs. active structures of class A GPCRs (Ring, 2013; Deupi, 2012; Kruse, 2013), L412<sup>5.51</sup> is highly conserved across class F GPCRs and also appears in a conformationally labile region of GPCRs. In class A GPCRs, residue 5.51 is one of a group of conserved hydrophobic and aromatic residues (3.40, 5.51, 6.44, 6.48) thought to constitute a “transmission switch” that rearranges when agonist binds (Venkatakrisnan, 2013; Deupi, 2011). Collectively, these CAMs bolster the notion Smo cycles through canonical GPCR inactive-active states.

Vismodegib is a Smo inhibitor that binds to the 7TM pocket (Figure 3.4) and has been approved for the treatment of advanced BCC. Resistance to Vismodegib usually appears within a few months, however (Chang, 2012). Cancers with active Hh signaling

are often driven by inactivating Ptc mutations, but resistance mutations often appear in Smo, the target of the drug. The Vismodegib resistance mutation originally found in medulloblastoma, D473H (Yauch, 2009), disrupts Vismodegib binding to Smo but does not result in Smo activation or loss of Smo regulation by physiological levels of Ptc. Additional drug resistance mutations in Smo were found in a mouse model of medulloblastoma where treatment with NVP-LDE225, a Smo 7TM antagonist, led to resistance mutations in Smo that predominantly localize to the 7TM binding pocket and result in phenotypes similar to D473H (Buonamici, 2010). Several unique Smo resistance mutations (W281L<sup>2,57</sup>, V321M<sup>3,32</sup>) were also recently found in BCC after treatment with Vismodegib (Brinkhuizen, 2014). W281L<sup>2,57</sup> localizes to the base of the 7TM binding pocket within 3.7 Å of the base of the LY2940680 ligand. V321M<sup>3,32</sup> is further buried at the base of the binding pocket and 5.8 Å from SANT1 at its closest point. It is not known whether these mutations function to disrupt binding of Vismodegib to Smo or to activate Smo, but its position in the Smo structure suggests that W281L is more likely to interfere with ligand binding than V321M. Given the rapid resistance to drugs targeting the Smo 7TM pocket, antagonists that bind the Smo CRD hold out the hope that drugs targeting the CRD may prove more effective or less susceptible to resistance when used either alone or in combination with compounds targeting the Smo 7TM pocket (Nedelcu, 2013; Nachtergaele, 2013).

Any discussion of the Smo 7TM and CRD regions naturally leads to questions concerning how these components interact and how their interplay affects the Smo C-terminal tail. Little is known about the structure of the Smo C-tail alone or with the Smo 7TM bundle, but its low complexity and high hydrophilicity suggest it doesn't adopt a

rigid globular structure. The Smo C-tail is phosphorylated in response to pathway activation although the identities of the kinases responsible for phosphorylation differ between vertebrates and invertebrates Jia, 2004 ; Chen, 2004). A conformational change of the *Drosophila* Smo C-tail has been proposed to stem from C-tail phosphorylation altering interactions between positively charged clusters of Arg residues and negatively charged clusters of Asp residues (Zhao, 2007), but the vertebrate Smo C-tail does not possess the Arg clusters. A C-tail conformational change in vertebrates has also been proposed, however (Chen, 2011).

### *Summary*

Multiple inputs—oxysterol binding to the CRD, small molecule binding to the 7TM pocket, and sterols within the cell membrane—are all capable of modulating Smo activity and presumably conformation. Sorting out what the endogenous inputs are, which of these inputs are important in specific instances, how multiple inputs are integrated, how best to exploit various ways of modulating Smo for anticancer therapies, and the role of Ptc in modulating these and perhaps other inputs present exciting challenges. Recent results have helped clarify the nature and sites of these inputs, however, and provided a framework for understanding how each of the parts fit together to achieve remarkable biological results.

The onset of biochemical characterization of the 7TM region and the CRD domain alone, including panels of inhibitory and activating small molecules for each domain stimulated efforts on our part to express, purify, characterize, and determine the three-dimensional structure of the CRD-7TM Smo.

## Methods

### *Expression of Smo in Sf9 Cells*

Smo variants for expression were designed and initial baculovirus directing their expression were generated by Ben Myers. The two main constructs discussed are 7TM Smo, which is identical to the construct crystallized in the Stevens lab (Wang, 2013). Abbreviated in this work as 7TM Smo, the construct used was bRIL-hSmo- $\Delta$ CRD- $\Delta$ CT. The apocytochrome b<sub>562</sub>RIL was fused to the N-terminus of the CRD domain. N-terminal affinity tags, a flag epitope followed by a 6X His sequence, were separated from bRIL by a TEV protease recognition site. The primary CRD-7TM Smo construct used in protein purification and crystallization attempts was bRIL-mSmo- $\Delta$ CT. The placement of bRIL and affinity tags was the same as in the 7TM Smo constructs. Further manipulations of the CRD-7TM Smo variants involved removal of bRIL from the N-terminus and insertion of T4 lysozyme (T4L0 at intracellular loop 3. Expression methods were identical for all variants,. Baculovirus at an MOI of 5 was used to infect Sf9 cells for 72 hours. Sf9 cells were harvested and flash frozen prior to protein purification.

### *Purification of mSmo constructs*

Protein purification followed methods described in (Wang, 2013) and illustrated in Figure 3.5. In brief, isolated membranes were washed three times and resuspended by dounce homogenization. Resuspended membranes were solubilized in 1%DDM/0.2%Cholesterol Hemisuccinate (CHS). TALON resin was used for initial protein capture and desalted TALON elution fractions were incubated with TEV protease for tag removal. Reincubation with Nickel NTA resin was used to remove His-tagged TEV protease and proteolyzed tags. Flow through from the Nickel NTA resin was



concentrated to 50 mg/ml for crystallization attempts. A portion of this flow through was used for analytical SEC on a Superose 6 column pre-equilibrated with 0.03%DDM, 0.006% CHS to assess protein monodispersity. For protein intended for use in crystallization trays, modulators (i.e. Smo agonists and/or antagonists likely to stabilize active or inactive states of Smo) were added at the point of solubilization and maintained throughout protein purification. For protein intended for thermal denaturation analysis, modulators were absent during prep and added prior to denaturation assay.

#### *Thermal denaturation of mSmo variants*

Thermal denaturation assays were performed using CPM (N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide) dye, a thiol-specific fluorochrome from Invitrogen. A 1:40 dilution of a 4 mg/ml solution of CPM dye (Invitrogen) in DMSO was incubated for 5 min at room temperature, protected from light. 20 µg of Smo was diluted to a final volume of 290 µL. After 5 min at room temperature, 10 µL of dilute dye was added to the Smo sample, which was mixed and transferred to a quartz fluorometer cuvette. The cuvette was transferred to a Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon) equipped with a Peltier sample cooler (F-3004) and heated at a rate of 2°C/min. Emission at 463 nm was monitored with an excitation wavelength of 387 nm from 20°C to 80°C. A melting temperature ( $T_m$ ) was determined by fitting the curve to a sigmoidal Boltzmann equation.

#### *Lipidic cubic phase crystallization*

Concentrated Smo protein was mixed with molten 90:10 monoolein:cholesterol at a ratio of 2:3 (w:w) using two Hamilton syringes connected by a syringe coupler. Mixing was performed by slowly mixing protein with lipid at a rate of 1 syringe stroke

per second. Successful formation of the lipid cubic phase (LCP) was noted by transition of the mixture from non-uniform cloudiness to the optically transparent lipidic cubic phase. The LCP mixture was then transferred to a 10 microliter Hamilton syringe and a repeating dispenser was used to aliquot 200 nanoliter drops onto glass sandwich crystallization plates. One microliter of the precipitant solution was overlayed on the LCP drop. Trays were sealed and incubated protected from light in a temperature controlled room.

## **Results**

7TM and CRD-7TM Smo proteins were expressed in Sf9 cells and purified using nickel affinity and size exclusion chromatography (SEC) (Figure 3.5, 3.6 and 3.7). Both the 7TM and CRD-7TM variants were monodispersed, eluting immediately prior to the 158kDa molecular weight standard (Figures 3.6B and 3.7B). Accounting for the DDM/CHS micelle, Smo likely elutes as a dimer compared to molecular weight standards; however, further evaluation using MALS or alternative SEC columns may better resolve the oligomeric state of Smo. The oligomeric state did not change with the addition of the CRD to the 7TM construct. Maximum protein yields for 7TM Smo were 300ug per liter of Sf9 cells. Yields for the 7TM construct were consistent across protein preparations; however, CRD-7TM yields fluctuated with an average of 150ug per liter. CRD-7TM Smo had variable amounts of CRD cleavage from the 7TM domain over the course of protein purification. This cleavage occurred differently when different modulators were added and is likely the source of inconsistent protein yields. Changes in

media formulation and baculovirus multiplicity of infection (MOI) were used to optimize protein yields.

The effect of small molecule modulators of CRD-7TM Smo were determined using thermal denaturation assays. The apo CRD-7TM Smo had a  $T_m$  of 54.5°C and its thermal denaturation curve was cooperative (Figure 3.8). The addition of a small molecule targeting either the Smo 7TM or Smo CRD domain did not affect the melting temperature when added individually. However, when a modulator of the 7TM and a modulator of the CRD were added in combination, the  $T_m$  of CRD-7TM Smo increased by 2.5°C. This change in  $T_m$  occurred whether 7TM modulator was an activator (Figure 3.8A) or inhibitor (Figure 3.8B). All crystal trays of CRD-7TM Smo included both a 7TM and CRD modulator in efforts to maximize protein stability.

Purified 7TM Smo and CRD-7TM Smo purified proteins, at concentrations of 50mg/ml, were used in lipidic cubic phase crystallization. Although I was able to reproduce the 7TM Smo crystals in the presence of LY2940680 in a published crystallization condition (Wang, 2013), no other crystals were obtained. Crystallization attempts included addition of a high affinity agonist, sSAG, with the 7TM Smo and a panel of 7TM agonists and antagonists along with 20(S)-OHC for the CRD-7TM Smo. The CRD-7TM Smo variant failed to give reasonable amounts of precipitate in any conditions tested. Further manipulations to CRD-7TM Smo, including alternative placement of the bRIL domain and addition of T4L to intracellular loop 3 were tried; however, the removal of the bRIL domain from the N-terminus of the CRD-7TM Smo resulted in apparent instability of the protein throughout purification, and these additional

CRD-7TM Smo proteins did not produce sufficient yields to be used in LCP crystallization.

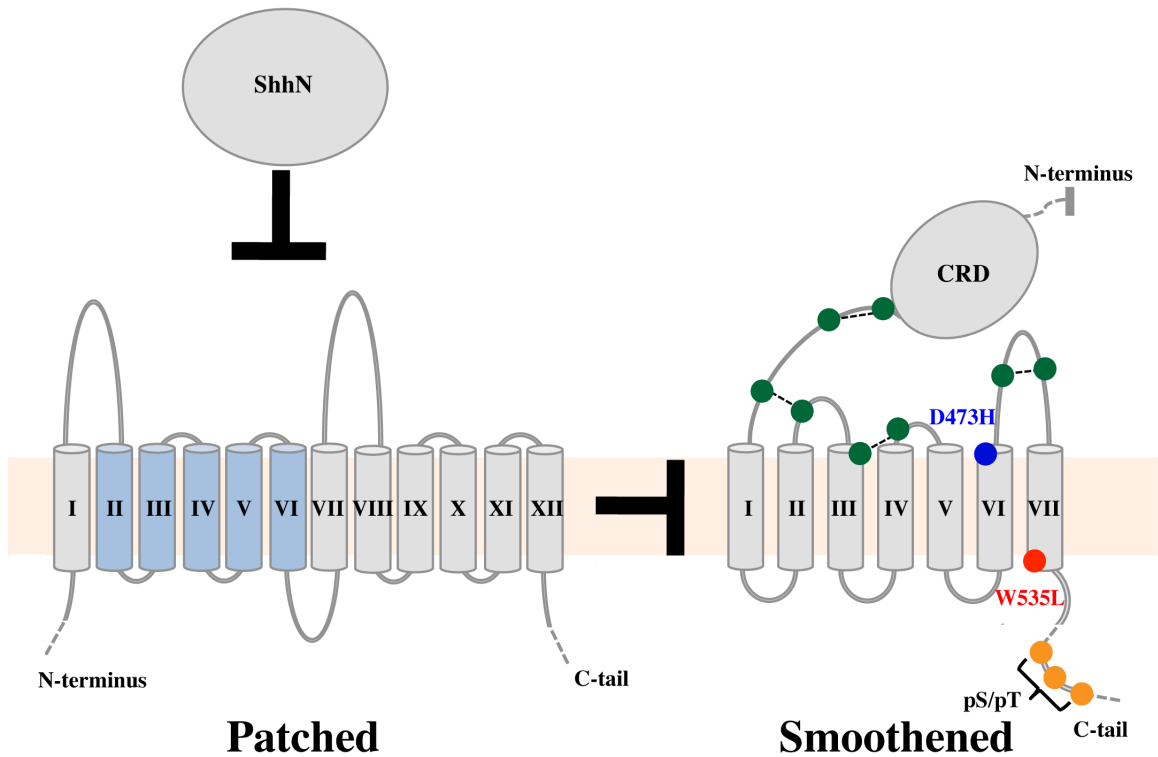
## **Discussion**

Efforts to crystallize 7TM Smo in the active state and CRD-7TM Smo in the presence of modulators to both domains proved unsuccessful. There are currently no structures of CRD-7TM Smo and while the 7TM Smo has been crystallized in the presence of one agonist, it failed to adopt a 7TM active state confirmation (Wang, 2014). Future efforts to capture a 7TM active state, if possible, may require the use of conformational specific antibodies/nanobodies. The CRD-7TM Smo variant presents additional difficulties due to the lack of protein precipitation in all crystallization conditions tried, which members of the Kobilka lab at Stanford indicated is unfortunately a sign that crystallization is extremely unlikely.

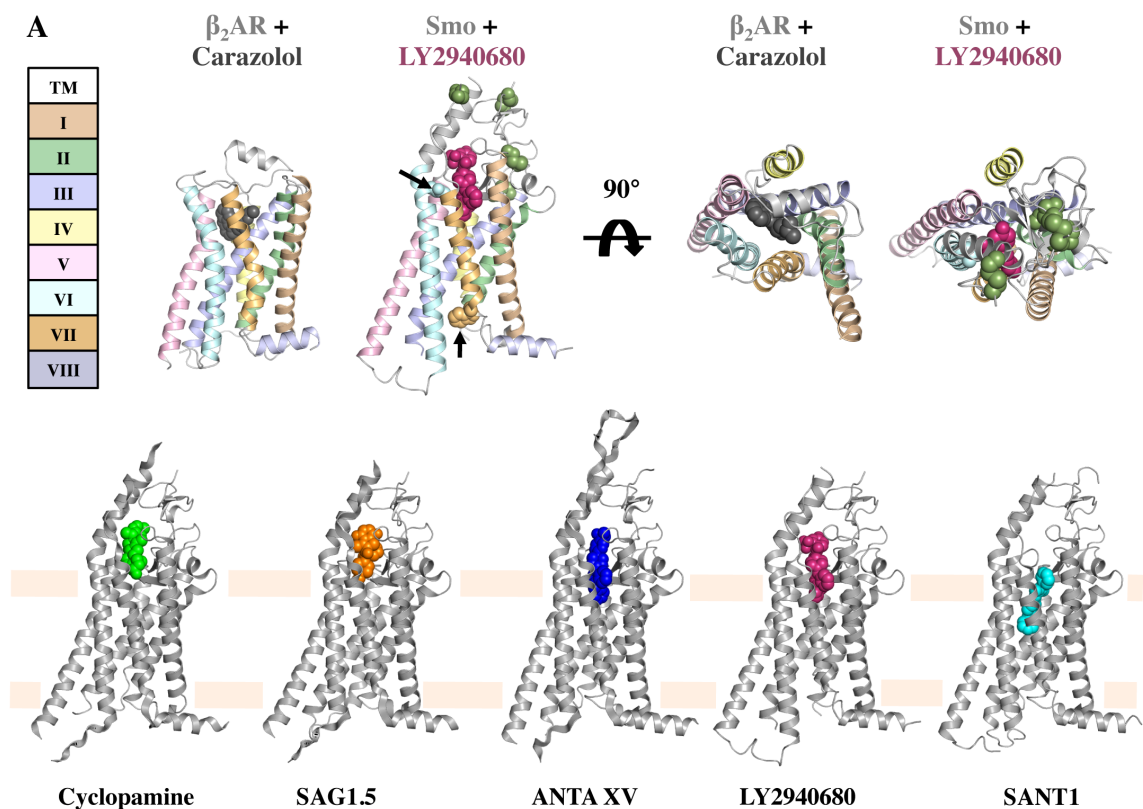
Nonetheless, multiple strategies remain to be tried in future attempts to crystallize CRD-7TM Smo. First, the CRD-7TM domain, with the removal of the bRIL domain from the N-terminus, was largely instable. Screening of Smo homologs from different species may prove useful. Indeed, the crystal structure of the CRD domain alone used the zebrafish Smo CRD. Conformational specific nanobodies that prevent CRD cleavage from the 7TM domain could also prove useful. Second, given the lack of precipitate in the most common lipidic cubic phase lipid, monoolein, additional LCP lipids may be tried. Alternatively, the CRD-7TM Smo construct may not lend itself to LCP crystallization and more traditional hanging drop crystal trays with Smo in detergent

may prove more fruitful. Finally, further analysis of modulators that stabilize the CRD-7<sup>TM</sup> Smo domain beyond the 2.5°C increase seen here may be helpful.

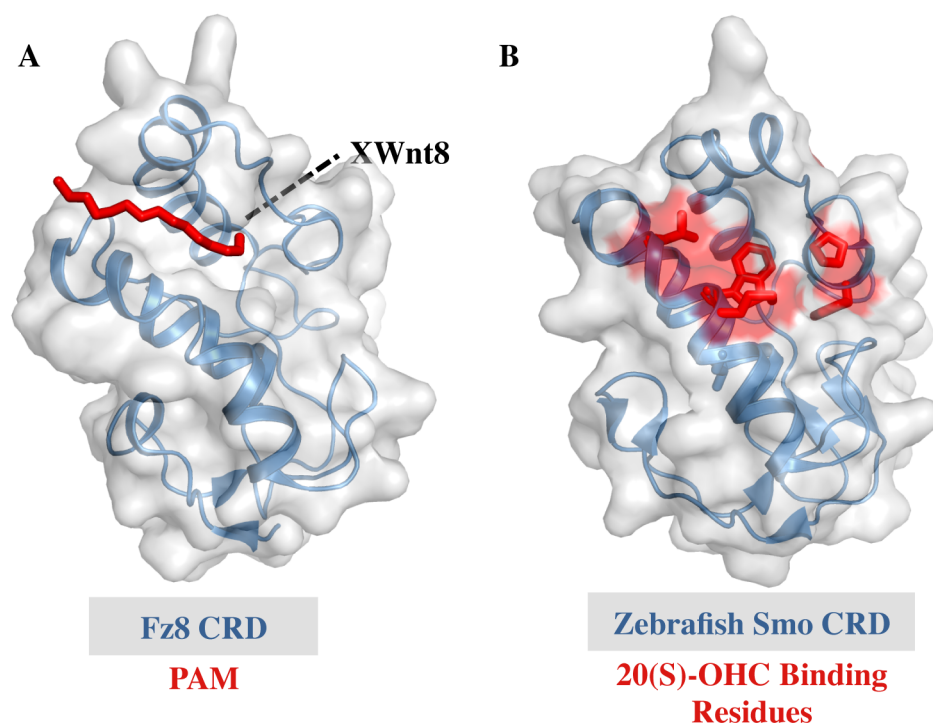
As our understanding of the biochemistry of the Smo CRD and of targeted Smo therapeutics increases, additional small molecules that stabilize Smo or aid in crystallization may be found. Ultimately, the current lack of understanding of how Ptc regulates Smo presents the most compelling question concerning Smo biochemistry and structure. The key to unlocking the CRD-7<sup>TM</sup> Smo structure may lie in the identification of the enigmatic regulator by which Ptc modulates Smo by function.



**Figure 3.1 Major transmembrane components of Hh signal reception and transduction.** Ptc (left) represses Smo (right) through an unknown, indirect mechanism. The interaction of Shh with Ptc relieves Ptc-mediated repression of Smo. The SSD of Ptc (TM II-TM VI) is colored blue. For Smo, the 8 cysteines mediating 4 disulfide bonds in the Smo ECLs are shown in green, the D473H Vismodegib resistance mutation in blue, the constitutively active W535L substitution in red, and pS/pT sites in the C-tail in orange.

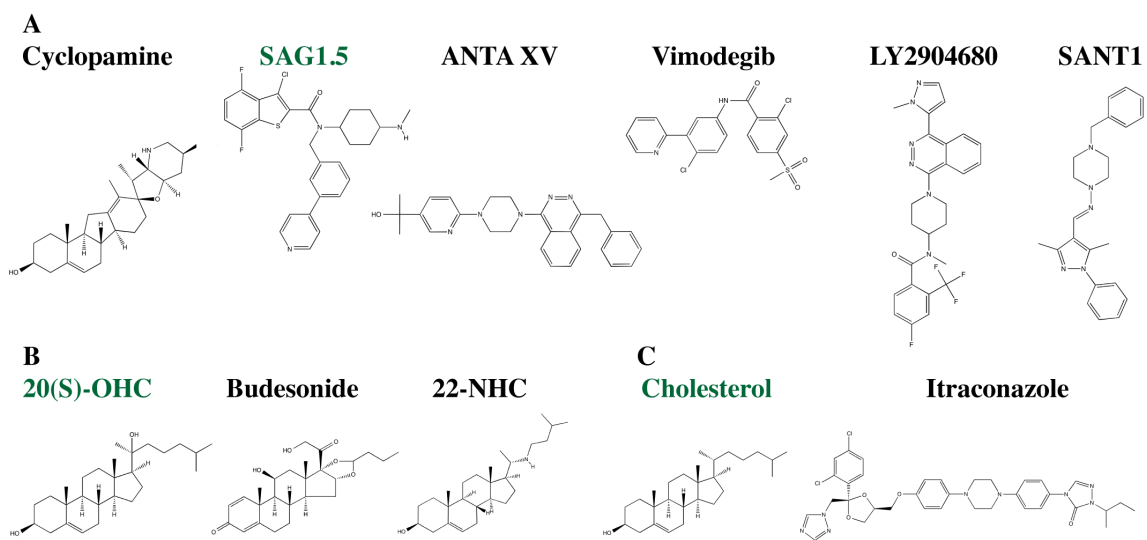


**Figure 3.2 Smo 7TM Region.** A.  $\beta_2\text{AR}$  in complex with Carazolol (PDB: 2RH1) and Smo in complex with LY2940680 (PDB: 4JKV) colored by GPCR helix number. Key Smo residues are shown in spheres (ECL disulfides, green; D473H, light blue; W535L, orange) B. All five Smo 7TM and ligand crystal structures (From left to right, PDB: 4O9R, 4QIN, 4QIM, 4JKV, 4N4W).

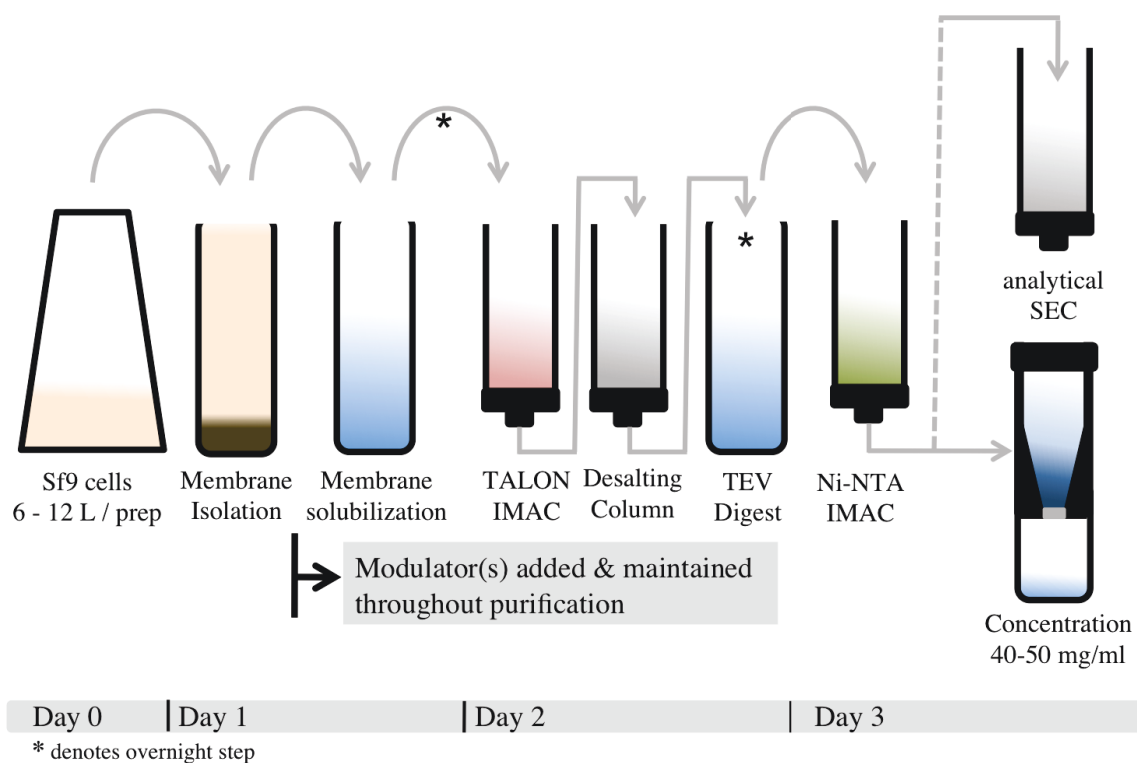


**Figure 3.3 Structure of class F GPCR CRDs.** (A) The structure of the mouse Frizzled CRD 8 (PDB: 4F0A) shown with the palmitoleic acid moiety (red) from XWnt8. The position of XWnt8 (structure not shown) is noted with a dashed black line. (B) The structure of the zebrafish Smoothened CRD (PDB: 4C79) with 20(S)-OHC binding residues shown in red.

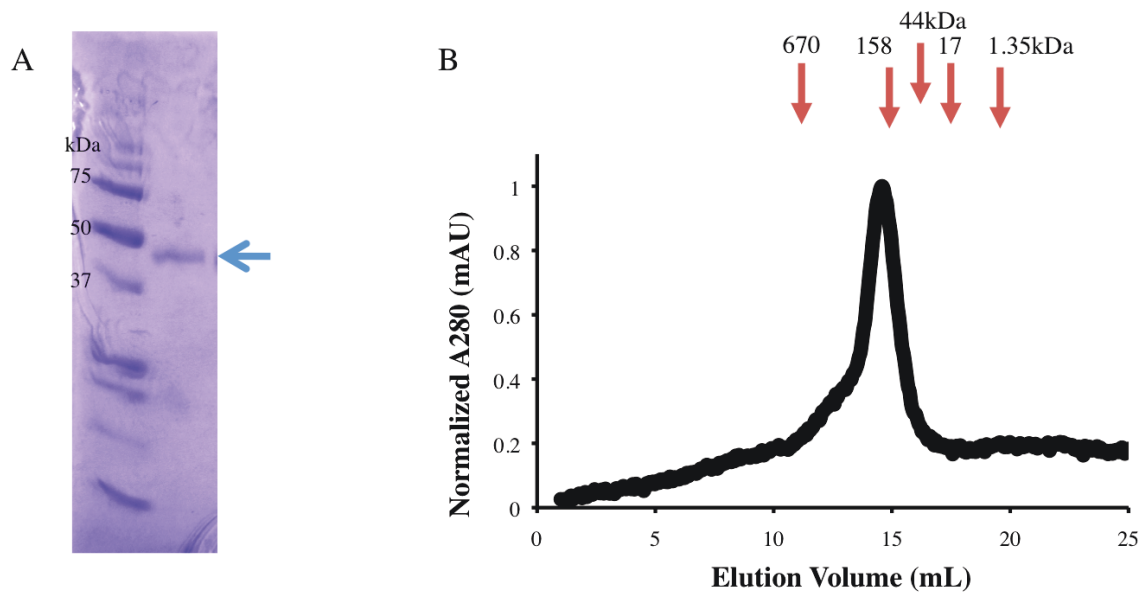




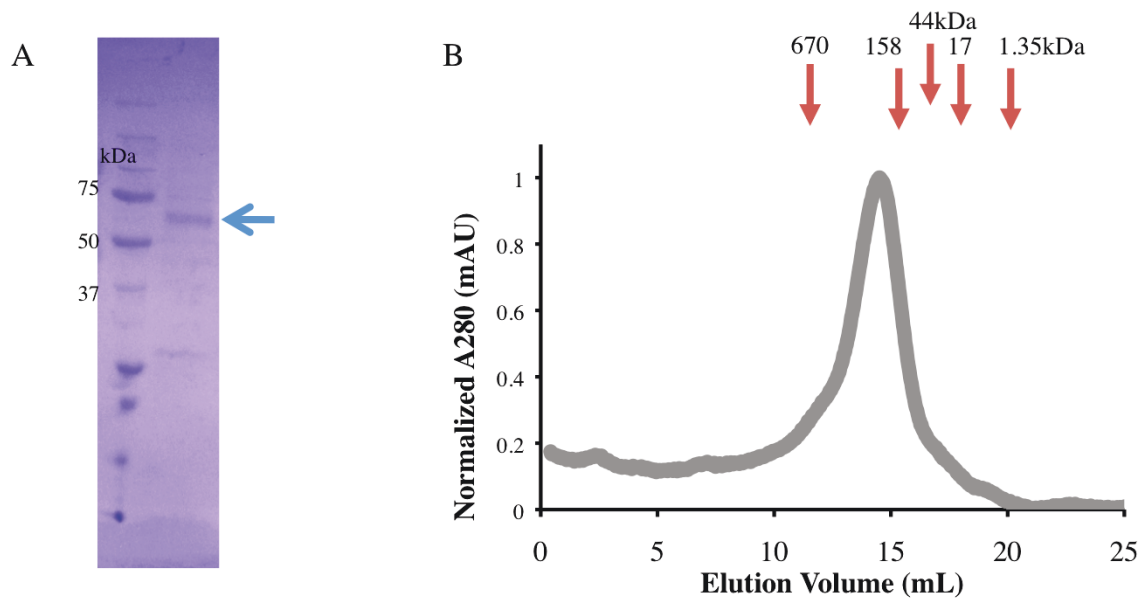
**Figure 3.4 Smoothened-interacting small molecules.** (A) 7TM targeting small molecules. (B) CRD targeting small molecules. C. Other Smo targeting small molecules. Activating small molecules are noted by green font.



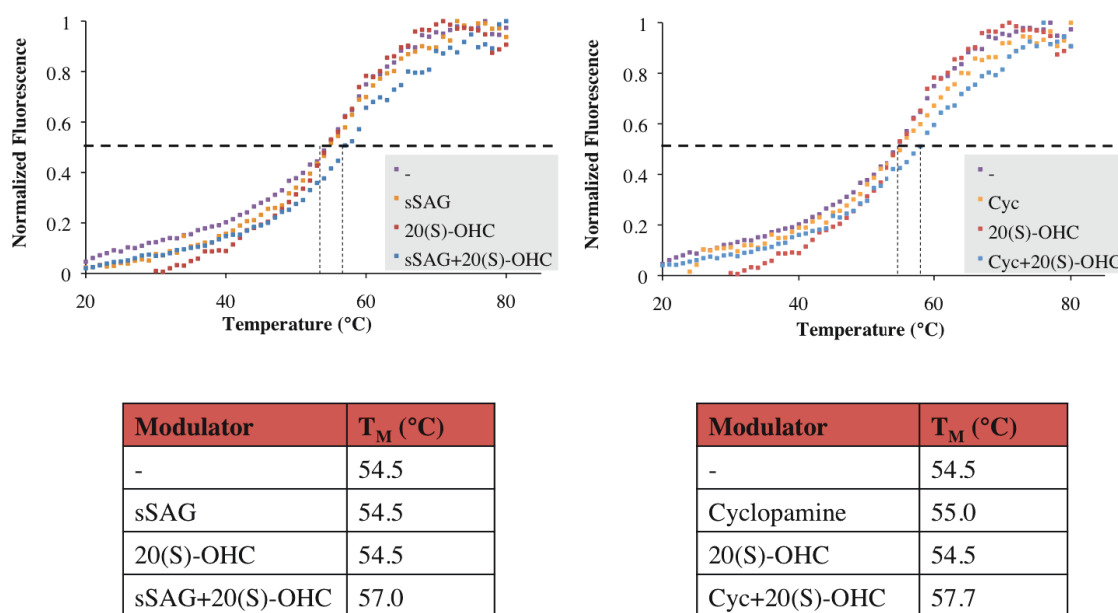
**Figure 3.5 Purification of CRD-containing Smo variants.** All Smo variants were purified from baculovirus infected Sf9 cells using nickel affinity chromatography followed by tag removal and size exclusion chromatography.



**Figure 3.6 Purification of mSmo-ΔCRD-ΔCT** (A) Purified 7TM Smo is shown on a Coomassie Brilliant Blue stained SDS-PAGE gel. (B) 7TM Smo is monodisperse on a Superose 6 column equilibrated in buffer containing 0.03%DDM/0.006%CHS.



**Figure 3.7 Purification of CRD-mSmo- $\Delta$ CT** (A) Purified CRD-7TM Smo is shown on a Coomassie stained SDS-PAGE gel. (B) CRD-7TM Smo is monodisperse on a Superose 6 column equilibrated in buffer containing 0.03%DDM/0.006%CHS.



**Figure 3.8 Thermal denaturation of purified Smo protein** Thermal denaturation curves of purified CRD-7TM Smo using a thiol specific dye (CPM) in the absence or presence of indicated activators or inhibitors. Melting temperatures (T<sub>m</sub>) were determined by fitting curves to the sigmoidal Boltzmann equation.

## **Chapter 4. Detergent-solubilized Patched Purified from Sf9 Cells Fails to Interact Strongly With Cognate Hedgehog or Ihog Homologs**

### **Acknowledgements**

This chapter is published in *Protein Expression and Purification* by Thomas E. Cleveland IV, Jacqueline McCabe and Daniel J. Leahy (Cleveland, 2014). My work focused on characterization of the Patched (Ptc) protein used in the pulldown assays. I screened our expressed and purified Ptc construct in a panel of detergents on size exclusive chromatography (Figure 4.2). I determined that membranes isolated from Sf9 cells expressing the mouse Ptc variant are capable of binding ShhN (Figure 4.1). I performed thermal denaturation analysis of Ptc to further characterize its behavior in detergent micelles (Figure 4.3). I led initial efforts to coexpress Ptc with ShhN with the goal of Ptc protein of improved quality. I performed all of the manuscript re-submission experiments required for final publication.

### **Abstract**

Patched (Ptc) is a twelve-pass transmembrane protein that functions as a receptor for the Hedgehog (Hh) family of morphogens. In addition to Ptc, several accessory proteins including the CDO/Ihog family of co-receptors are necessary for proper Hh signaling. Structures of Hh proteins bound to members of the CDO/Ihog family are known, but the nature of the full Hh receptor complex is not well understood. We have expressed the *Drosophila* Patched and Mouse Patched-1 proteins in Sf9 cells and find that Sonic Hedgehog will bind to Mouse Patched-1 in isolated Sf9 cell membranes but that purified, detergent-solubilized Ptc proteins do not interact strongly with cognate Hh

and CDO/Ihog homologs. These results may reflect a nonnative conformation of detergent-solubilized Ptc or that an additional factor or factors lost during purification are required for high-affinity Ptc binding to Hh.

## **Introduction**

The Hedgehog (Hh) signaling pathway mediates key tissue patterning events during animal development, and abnormal pathway activity is associated with several cancers (Johnson, ., 1996; Goodrich, 1997). Hh proteins are secreted morphogens that specify cell fates in neighboring tissues in a concentration-dependent manner (Riddle, Johnson, Laufer, & Tabin, 1993; Heemskerk & Dinardo, 1994). The twelve-pass transmembrane protein Patched (Ptc) has been identified as a key Hh receptor in genetic (Hidalgo, 1990; Ingham, 1991; Forbes, 1993; Marigo, 1996) and cell-based binding studies (Stone, 1996; Fuse, 1999). In the absence of Hh, Ptc constitutively inhibits the activity of Smoothened (Smo) (Ingham, 1991; Taipale, 2002), a seven-pass transmembrane protein. The mechanism of this inhibition is unknown but does not appear to involve a direct interaction between Ptc and Smo (Taipale, 2002). In the presence of Hh, this inhibition is relieved and the pathway is activated. For recent reviews, see (Beachy, 2010; Ingham, 2012; Briscoe & Théron, 2013; Ryan & Chiang, 2012).

Many additional proteins modulate Hh pathway activity, but their presence and activity are not always conserved across phyla. For instance, Gas1 positively regulates Hh signaling in mammals (Allen, 2011), but no Gas1 homolog exists in the fruit fly. Hhip, a cell surface protein that acts as a negative regulator by binding and sequestering Hh proteins in vertebrates (Chuang & McMahon, 1999) also has no apparent ortholog in

the fly. The mammalian proteins CDO and BOC are orthologous to the fly proteins Ihog and Boi and each binds its cognate Hh protein, but the manner and co-factor dependence of Hh binding by fly and mammalian orthologs is not conserved (McLellan, 2008).

Despite the central importance of the Hh signaling pathway in animal development and the identification of many key pathway components, little is known about the molecular details connecting these components. Ptc is presumed to control Smo activity by transporting a small sterol-like molecule, a hypothesis based on an array of circumstantial evidence: the homology of Ptc to proton antiporters in the RND superfamily; the presence in Ptc of a sterol-sensing domain, which in other eukaryotic homologs is related to cholesterol trafficking; the indirect inhibition of Smo by Ptc, which suggests an intermediate (Taipale, 2002); and recent structural and biochemical studies of Smo showing that sterols bind Smo and modulate its activity (Wang, 2013; Nedelcu, 2013; Myers, 2013). Nevertheless, ligand transport by Ptc has not been conclusively demonstrated, nor has a physiological ligand for Smo been identified. More generally, how Hh proteins modulate the activity of Ptc is not known, and no functional assay for purified Ptc has been established.

A direct interaction between Hh and Ptc is clearly the simplest and most likely interpretation for high-affinity Hh binding to Ptc-expressing cells and Hh modulation of Ptc activity (Marigo, 1996; Fuse, 1999). Assays measuring binding to the cell surface leave open the possibility that other cellular factors could be involved, however. For example, most cell-based binding studies predated knowledge of the importance of accessory proteins Gas1, CDO/Ihog and BOC/Boi for mediating interactions between Hh and the cell surface. These proteins are essential for both fly (Zheng, 2010) and



mammalian Hh signaling (Allen, 2011; Izzi, 2011). Members of the CDO family of proteins bind Hh proteins directly with low micromolar affinities (McLellan, 2006; McLellan, 2008; Kavran, 2010), whereas the affinity of mammalian Sonic Hedgehog (Shh) for binding to the surface of Ptc-expressing cells is in the low nanomolar range (Marigo, 1996; Fuse, 1999). This difference in affinity, as well as the fact that high-affinity cell-surface binding is dependent on the presence of Ptc, would seem to indicate that Ptc itself directly binds Shh but that the affinity for the binary Shh-Ptc interaction may be weaker than the affinity of Hh for Ptc-expressing cells. The role of the CDO/Ihog family of proteins appears to be to function as co-receptors that enhance binding affinity. In this case, a ternary complex between Ptc, Hh and members of the CDO/Ihog family may represent the initial signaling complex at the cell surface. Evidence for such ternary complexes has been found for *Drosophila* Ptc and Ihog (Zheng, Mann, Sever, & Beachy, 2010), but the evidence in the mammalian pathway is contradictory. Although CDO, BOC or Gas1 appear to be required along with Ptc for normal signaling in mammals (Allen, 2011; Izzi, 2011), the addition of the soluble Hh-binding domain of CDO (CDOFn3) actually competes for ShhN binding to Ptc on the cell surface (McLellan, 2008). This observation suggests that the binding surfaces for Ptc and CDO on ShhN may overlap. This binding competition has been rationalized with the positive role of both Ptc and CDO in Hh signaling by the observation that physiological Hh is found in multivalent particles (Chen, 2004), allowing simultaneous Ptc and CDO binding. Multivalency is not required for high-affinity binding of Shh to Ptc on cells, however, as monomeric ShhN expressed in *E. coli* binds Ptc-expressing cells with high affinity (Marigo, 1996; Fuse, 1999). Hh proteins also bind sulfated glycosaminoglycans with

affinities that can be in the micromolar range, which also likely contributes to high-affinity interactions with the cell surface.

A major barrier to understanding Ptc activity and the nature of its interactions with Hh pathway components has been the difficulty of isolating functional forms of Ptc or Ptc fragments. Most information on the reported molecular mechanisms and binding partners of Ptc has been obtained indirectly, using cell, tissue or whole animal-based studies. We therefore undertook to express and purify mouse and *Drosophila* Ptc proteins with intact transmembrane and extracellular regions for binding studies with Hh and other Hh pathway components. Sonic Hedgehog will bind to mouse Ptc in isolated Sf9 cell membranes, but we surprisingly find that Ptc proteins extracted and purified in the presence of detergents do not interact with soluble, cognate Hh or CDO/Ihog homologs with high affinity as either binary or ternary complexes.

## **Methods**

### *Materials*

All chemicals were purchased from Sigma unless otherwise noted. Detergents were purchased from Affymetrix and included n-dodecyl- $\beta$ -D-maltopyranoside (DDM), n-dodecylphosphocholine (fos-choline 12, FC-12), and 2,2-didecylpropane-1,3-bis- $\beta$ -D-maltopyranoside (lauryl maltose neopentyl glycol, LMNG). Low molecular weight heparin (LMWH, average molecular weight 3000 Da, sodium salt, from porcine sources) was purchased from P212121.com. Heparin Decasaccharide (DH) was purchased from Neoparin Inc. Antibodies for western blotting were mouse  $\alpha$ -Myc monoclonal (9E10), which was isolated from hybridoma growth medium; rabbit  $\alpha$ -Ihog polyclonal and mouse

$\alpha$ -Hh monoclonal antibodies, which were gifts from P. Beachy; and appropriate HRP-conjugated secondary antibodies.

### *Buffers*

Ni<sup>2+</sup> Binding Buffer consisted of 35 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 15 mM imidazole adjusted to pH 8.0 with NaOH at 25°C. Ni<sup>2+</sup> Wash Buffer consisted of 10 mM Tris base, 10 mM Tris HCl, and 300 mM NaCl. Ni<sup>2+</sup> Elution Buffer was Ni<sup>2+</sup> Wash Buffer with 250 mM imidazole. Strep Wash Buffer was 20 mM MOPS, 10 mM NaOH, and 200 mM NaCl. Strep Elution Buffer was Strep Wash Buffer with 2.5 mM desthiobiotin. Pull-Down Buffer was Strep Wash Buffer containing 0.01% LMNG and 0.2 mM TCEP. CD Buffer was 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaF, titrated to pH 7.2 with NaOH. CPM Thermal Stability Buffer was 20 mM HEPES pH 7.5, 200 mM NaCl, 0.02% LMNG.

### *Proteins and expression vectors*

IhogFn1, IhogFn2, IhogFn12 and IhogFn12 $\Delta$ H are the first, second, or both Type III Fibronectin (FnIII) domains from Ihog, respectively, with  $\Delta$ H referring to an Ihog surface mutant with reduced heparin binding (McLellan, 2006). BOCFn3, BOCFn23 and BOCFn13 are FnIII domains 3, 2–3, and 1–3 respectively from BOC. These BOC and Ihog fragments were expressed as His-Myc-TEV- (HMT-) fusion proteins using the vector pT7HMT (Geisbrecht, 2006) HhN, ShhN and ShhN-SC (The “Surface C” mutant of ShhN (Fuse, 1999), which is deficient in binding to Ptc-expressing cells) were subcloned into a modified version of pMAL-c2x as described (McLellan, 2008). ShhFL and HhFL are the full-length Mouse and *Drosophila* Hh proteins including native signal sequences and C-terminal self-splicing domains. IhogFn12TM consisted of the native

Ihog signal sequence followed by the first and second FnIII domains and the transmembrane region, but with the intracellular region truncated. Ski and Hhat are the entire native Hh acyltransferases from *Drosophila* and Mouse, respectively. MmHhip consisted of the  $\beta$ -propeller and following two EGF domains of Mouse Hhip. *Drosophila* Patched and mouse Patched-1 proteins are the tagged Mouse and *Drosophila* Ptc proteins with their C-termini truncated immediately after the final predicted transmembrane helix (DmPtcT1 and MmPtcT1, respectively).

DmPtcT1 and MmPtcT1 proteins were expressed fused to a concatenated series of N-terminal tags including the Streptavidin Binding Peptide (SBP) (Keefe, 2001) an HRV3C protease site, a 6 $\times$ His tag, a Myc tag, and a TEV protease site. Tags and Ptc proteins were cloned into the transfer vector pFastBac1 (Invitrogen). For co-expression of ShhFL with Hhat, HhFL with Ski, and DmPtcT1 with IhogFn12TM, transfer vectors were constructed from pFastBacDual (Invitrogen). Hhip was cloned into a modified pFastBac1 vector containing the Honeybee Mellitin (HBM) signal sequence to target for secretion, followed by 8 $\times$ His, SBP, Myc and TEV sequences.

#### *Production of recombinant baculoviruses*

Recombinant bacmids and baculoviruses for insect cell expression were constructed using the Bac-to-Bac system (Invitrogen) following manufacturer's instructions. After transfecting Sf9 cells with bacmid DNA, the secreted virus (designated P1) was amplified two more times (P2 and P3 viruses) following manufacturer's instructions. P3 virus was used for protein production.

*Bacterial expression and purification of HhN, IhogFn1, IhogFn2, IhogFn12, IhogFn12 $\Delta$ H, ShhN, ShhN-SC, BOCFn3, BOCFn23 and BOCFn13*

All expression plasmids were transformed into BL21 and plated on LB with appropriate antibiotics. Single colonies were picked, grown overnight, and used to inoculate TB in baffled flasks at 225 RPM and 37°C. Bacteria were grown to an optical density (600 nm) of around 0.8, at which point the incubator temperature was lowered to 16°C and the cultures allowed to shake for an additional hour. IPTG was then added to a final concentration of 0.5 mM, followed by expression for 24 hr. Bacteria were harvested by centrifugation and cell pellets stored at -20°C or -80°C.

Individual bacterial pellets were resuspended in ice-cold Ni<sup>2+</sup> Binding Buffer containing 0.1 µL/mL Benzonase Nuclease and 1 mM PMSF. Bacteria were then lysed using a French press, and extracts cleared by centrifugation and syringe filtration. All proteins were initially purified using 5 mL HisTrap columns (GE Healthcare) charged with Ni<sup>2+</sup>. Where appropriate, tags were then removed by addition of His-tagged TEV or HRV3C and overnight incubation at 4°C. HhN, IhogFn12, IhogFn12ΔH, IhogFn1, and ShhN were then further purified by cation exchange chromatography as described (McLellan, 2006; McLellan, Zheng, Hauk, Ghirlando, Beachy, & Leahy, 2008). All proteins were desalted if necessary, then separated from tags and proteases by another round of HisTrap purification. All proteins were purified by a final size exclusion chromatography (SEC) step in Pull-Down Buffer, except for BOCFn13 for which the concentration of NaCl was 500 mM.

#### *Insect expression and purification of Hhip, DmPtc1 and MmPtc1*

Sf9 and High Five cells were adapted and grown in suspension using a modification of the serum-free medium ISFM (Inlow, Shauger, & Maiorella, 1989) containing 10 g/L glucose instead of 2.5 g/L. Sf9 or High Five cells were grown or split to a concentration

of  $2.0 \times 10^6$  cells/mL in air-sparged 1 L spinner flasks before infection with 20 mL of P3 baculovirus (for simultaneous infection with multiple viruses, 20 mL of each virus were added).

Hhip was expressed in High Five cells. After infection, expression was allowed to proceed for 72 hr before removing the cells by centrifugation. The conditioned medium was then adjusted to pH 8.0 with NaOH, resulting in a heavy precipitate of calcium phosphate that was removed by centrifugation and filtration. Up to 6 L of pH-adjusted, clarified medium was flowed by gravity through 15 mL of “Complete His” resin (Roche) at room temperature overnight, which was then washed and eluted using manufacturer’s recommended conditions. Eluted protein was further purified using a 5 mL StrepTrap column (GE Healthcare), and tags were cleaved by addition of TEV protease overnight. Cut tags, uncleaved fusion protein and TEV protease were then removed using a 5 mL HisTrap column, and the remaining Hhip was further purified by SEC.

MmPtcT1 and DmPtcT1 were expressed in Sf9 cells. After infection, expression was allowed to proceed for 72 hr before harvesting cells by centrifugation. If not used immediately, cell pellets were stored by washing with PBS containing 10% glycerol, followed by snap freezing in liquid nitrogen and storage at  $-80^{\circ}\text{C}$ . Cells were resuspended to a final volume of 30 mL per L of culture with  $\text{Ni}^{2+}$  Bind Buffer containing 0.5 mM TCEP, 0.1  $\mu\text{L/mL}$  Benzonase and one Complete Protease Inhibitor tablet, EDTA-free (Roche). Cells were lysed by French press, and nuclei and large debris removed by centrifugation at 10,000 rpm for 10 min in an SA-600 rotor. The supernatant was then transferred to a thin-walled ultracentrifuge tube and additional buffer added up to the maximum tube volume, around 40 mL, and centrifuged in an sw-28 rotor at 28,000

rpm for 1 hr to pellet membranes. Isolated membranes were weighed, added to 10 mL of  $\text{Ni}^{2+}$  Bind buffer per gram of membrane pellet, and resuspended using a Dounce homogenizer. 1 mL of 10% FC-12 per gram of membrane pellet was then added to the resuspended membranes, which were allowed to solubilize for 30 minutes on ice. Finally, detergent-solubilized protein was clarified with an additional ultracentrifugation spin.

Solubilized Ptc was applied to a 5 mL HisTrap column, washed with 50 mL of Nickel Wash Buffer supplemented with 0.2 mM TCEP and 0.02% LMNG, and eluted with Nickel Elution Buffer supplemented with 0.2 mM TCEP and 0.02% LMNG. A small amount of the reducing agent TCEP was added to prevent unwanted disulfide formation between exposed intracellular cysteines. At this point, purified Ptc could be incubated in batch with a Strep-Tactin bead slurry for use in pulldown assays. Alternatively, to obtain higher concentrations and purity for Circular Dichroism (CD), analytical SEC, or for pulldown by ShhN- or HhN-coupled beads, an additional round of purification was performed using a 5 mL StrepTrap HP column. Purified Ptc from the nickel column was applied to the StrepTrap column, washed with Strep Bind Buffer supplemented with 0.2 mM TCEP and 0.01% LMNG, and eluted with Strep Elution Buffer containing 0.2 mM TCEP and 0.01% LMNG. For Ptc proteins intended for CD, the Strep-Tactin purification buffers were substituted with CD Buffer containing the same amounts of TCEP, LMNG and desthiobiotin. For Ptc proteins intended for 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM) thermal melts, Ptc was purified as described above with the following modifications: 2 mg/ml iodoacetamide was added at membrane solubilization; all subsequent buffers did not include TCEP; and protein from Nickel column elution was

desalted using a 5mL HiTrap desalting column equilibrated in CPM Thermal Stability Buffer. Freshly purified protein was used the same day for CPM thermal melt assays.

*Size exclusion chromatography of detergent-exchanged Ptc*

For each detergent screened, a membrane pellet expressing MmPtcT1 from 125 ml Sf9 culture was used. Membrane pellets were solubilized in 1% fos-choline 12 as described in Methods. Solubilized membranes were loaded onto a 1 mL HiTrap Nickel column. The column was washed with 20 column volumes of buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 50 mM imidazole, 0.2 mM TCEP and 0.1% fos-choline 12. Bound protein was eluted in 10 column volumes of elution buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole, 0.2 mM TCEP and the greater of  $2 \times \text{C.M.C.}$  or 0.05% (w/v) of the detergent of interest. Nickel elutions in the detergent of interest remained at 4°C overnight and the next day were injected onto a Superose6 column. The Superose6 column was pre-equilibrated in 20 mM Tris pH 8.0, 300 mM NaCl, 0.2 mM TCEP and the detergent of interest at the indicated concentration (w/v) as follows. Detergents screened included octyl maltoside (1.78%), dodecyl maltoside (0.05%), decyl maltoside (0.345%), Cymal 4, Cymal 5 (0.24%), Cymal 6 (0.06%), Cymal 7 (0.05%), octyl glucose neopentyl glycol (0.12%), decyl maltose neopentyl glycol (0.05%), and lauryl dimethylamine N-oxide (0.05%).

*Immunostaining and confocal microscopy of Patched-expressing Sf9 cells*

Sf9 cells were plated on gelatinized glass coverslips in 6-well plates and simultaneously infected with the indicated viruses. Expression was allowed to proceed for 24-36 hours. All steps were performed at room temperature unless otherwise noted. Cells were fixed by adding to the growth medium an equal volume of PFA (4%



paraformaldehyde in PBS) for 5 min, followed by aspiration and addition of undiluted PFA for another 5 min. Cells were then permeablized with 0.5% Tween-20 in PBS for 15 min, and blocked with PBST (PBS with 0.1% Tween-20) containing 1% BSA for 30 min. Staining was performed with  $\alpha$ -Myc (1:200 in blocking solution) at 4°C overnight. Staining solution was removed and cells washed 3 times with PBST, incubating for 5 minutes each wash. Finally, cells were incubated in the dark for 1 hour with fluorescent secondary antibody solution (Cy3 Goat  $\alpha$ -Mouse, 1:1000 in blocking solution), washed 3 times with PBST, and coverslips mounted on glass slides in aqueous mounting medium for confocal microscopy.

#### *Cell-free binding assay*

Cell-free binding assay was performed as in (Zheng, 2010). In brief, isolated membrane pellets from uninfected Sf9 cells or cells infected with MmPtcT1 baculovirus were resuspended in ShhN-Renilla conditioned medium. After 1 hr incubation, membranes were reisolated by centrifugation, washed three times with ice cold PBS, and solubilized in 250 $\mu$ L passive lysis buffer (Promega) and 5 $\mu$ L of lysate used to measure Renilla luciferase activity.

#### *Ptc characterization*

The behavior of Ptc proteins was analyzed by CD, CPM fluorescence thermal stability assays, and by analytical SEC using a Superose6 column. For CD analysis, spectra were collected at 4°C between 185 nm and 280 nm, using Ptc proteins in a cuvette with 1 mm pathlength. Thermal unfolding analysis was also performed by following the CD signal at 208 nm while increasing the temperature from 4°C to 95°C. CPM fluorescence thermal stability assays were performed as in (Alexandrov, Mileni,

Chien, Hanson, & Stevens, 2008). In brief, a 1:40 dilution of a 4 mg/ml solution of CPM dye (Invitrogen) in DMSO was incubated in CPM Thermal Melt Buffer for 5 min at room temperature, protected from light. Ptc protein (10  $\mu$ g) was diluted in CPM Thermal Melt Buffer to a final volume of 290  $\mu$ L. After 5 min at room temperature, 10  $\mu$ L of dilute dye was added and the sample was mixed and transferred to a quartz fluorometer cuvette. The cuvette was transferred to a Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon) equipped with a Peltier sample cooler (F-3004) and heated at a rate of 2°C/min. Emission at 463 nm was monitored with an excitation wavelength of 387 nm from 20°C to 80°C. For SEC analysis, Ptc was injected onto a Superose6 column equilibrated in Pulldown Buffer. Its elution profile was followed by UV absorption at 280 nm, and apparent molecular weights were estimated by comparison to Gel Filtration Standards (Bio-Rad).

#### *Pulldown Assays*

Bead-based pulldown assays were used to screen for interactions between purified, detergent-solubilized Ptc and other proteins. All steps were carried out on ice or at 4°C unless otherwise noted. Immediately after performing the nickel purification, Ptc resin was prepared by coupling the SBP-tagged protein to Strep-Tactin resin (IBA). An excess of Ptc protein solution (typically the elution peak from an entire 1 L prep) was added to enough Strep-Tactin resin to give 10  $\mu$ L of Ptc resin for each desired pulldown condition; typically, this resulted in 2–5  $\mu$ g of coupled Ptc protein per condition. Ptc was allowed to couple to completion for 2 hr. Resin was then pelleted and resuspended in Pulldown Buffer, divided into separate tubes, and pelleted, such that 10  $\mu$ L of packed Ptc resin were obtained in each tube. Pulldowns were set up in a total volume of 50  $\mu$ L using the reaction compositions specified in Figure 4.4 and 4.4. Generally, pulldown conditions

contained 10  $\mu$ M of each soluble protein, and around 0.75  $\mu$ M of Ptc. These binding reactions were incubated for 1 hr. Ptc beads were then washed 4 times with 1 mL of Pulldown Buffer, and Ptc was eluted by adding 20  $\mu$ L of Pulldown Buffer supplemented with 2.5 mM biotin. The presence of proteins that bound and eluted with Ptc was then analyzed using Coomassie-stained SDS-PAGE gels and Western blots.

#### *Co-purification of Ptc with other proteins*

Ptc proteins were co-expressed with potential partners by infecting Sf9 cells with the following viruses or combinations of viruses: MmPtcT1 virus with ShhFL/Hhat dual expression virus; DmPtcT1/IhogFn12TM dual expression virus; DmPtcT1 virus with HhFL/Ski dual expression virus; and DmPtcT1/IhogFn12TM dual expression virus with HhFL/Ski dual expression virus. In each case, the Ptc protein was fused with N-terminal SbpPHMT tags, and its potential binding partners were untagged. Sequential nickel, Strep-Tactin, and Superose6 chromatography steps were performed as outlined above, and the presence of co-purifying complexes evaluated by Coomassie-stained SDS-PAGE and Western blot.

## **Results**

#### *Ptc Purification and Characterization*

Both *Drosophila* and Mouse Ptc expressed at levels of 0.5 – 3 mg of protein per liter of Sf9 cell culture. Confocal imaging of infected Sf9 cells expressing Ptc proteins indicated that Ptc was present at the cell surface, and membranes isolated from MmPtcT1-expressing cells bound Shh (Figure 4.1). Detergent solubilization trials indicated, however, that Ptc proteins were only solubilized well by the fos-choline class

of detergents, and FC-12 was selected for routine extraction. Solubilized Ptc proteins were purified by a combination of Nickel and Strep-Tactin affinity chromatography (Figure 4.3A). Yields of purified Ptc ranged from 20–50% of total expressed protein as judged by Coomassie-stained gels and comparison to BSA standards. Losses resulted from incomplete solubilization, incomplete capture, or incomplete elution from Nickel and were dependent on flow rates and the amount of protein present initially. Although most detergents do not solubilize Ptc proteins efficiently, we found that once Ptc had been solubilized, FC-12 could be exchanged for a variety of other detergents during the initial HisTrap purification. We therefore evaluated the behavior of Ptc proteins by SEC after exchange into different detergents (Figure 4.2). SEC behavior was qualitatively similar for most detergents tested, with a species eluting as a small oligomer, potentially consistent with a trimer (Lu, 2006) (Fig. 4.3B, arrow 1); a higher-order oligomer (arrow 2); and a small proportion in the void volume. Although the SEC behavior of Ptc after purification is similar in many detergents, Ptc in LMNG was less prone to precipitation over several days of storage. All binding studies were therefore carried out with Ptc exchanged into LMNG.

Characterization of purified Ptc proteins by CD spectroscopy revealed spectra typical of proteins with high alpha-helical content (Figure 4.4C). The secondary structure content of DmPtcT1 calculated from its CD spectrum was 47% helix and 13% sheet, similar to the values of 44% and 6% predicted from the primary sequence by PSIPRED (McGuffin, 2000) CD at 208 nm was used to follow unfolding of DmPtcT1, giving an unfolding midpoint of 58°C. We also followed Ptc unfolding using CPM dye, a measure that is more closely related to tertiary rather than secondary structure. By this method, unfolding

midpoints were 36°C and 37°C for MmPtcT1 and DmPtcT1, respectively (Fig. 4.2D).

Deglycosylation of purified MmPtcT1 with PNGaseF also resulted in a small but visible band shift on Coomassie-stained SDS-PAGE, consistent with the molecular weight of several glycans; MmPtcT1 contains 6 consensus N-linked sites, with total expected glycan mass of 6-7 kDa, assuming pauci-mannose glycosylation profiles typical of Sf9 cells (Aeed & Elhammer, 1994). These results are consistent with proper processing and membrane localization of expressed Ptc proteins.

#### *Pulldown of potential binding partners by MmPtcT1*

MmPtcT1-coupled resin was tested for its ability to pull down soluble ShhN, ShhN-SC, HMT-BOCFn3, HMT-BOCFn23 and HMT-BOCFn13 individually as well as in ternary combinations (Figure 4.4). 1 mM  $\text{Ca}^{2+}$  was included in most conditions as  $\text{Ca}^{2+}$  is known to be required for ShhN binding to BOCFn3. 200  $\mu\text{M}$  LMWH was included in one condition with BOCFn13 and ShhN, since both are known to be heparin-binding proteins. We also included conditions with 5 mM EGTA, a  $\text{Ca}^{2+}$  chelator, as a control to eliminate ShhN:BOCFn13 binding. Finally, we included Hhip in certain conditions to test its ability to block potential binding of ShhN to MmPtcT1. Strep-Tactin resin without conjugated MmPtcT1 (Figure 4.4) or with the tagged, irrelevant membrane protein AcrB conjugated was used as a control for nonspecific binding. We used Coomassie-stained SDS-PAGE to detect bound proteins, with Western blots also performed in some cases to detect weakly-bound proteins and confirm the identities of Coomassie-stained protein bands.

ShhN did not bind to detectably to MmPtcT1 beads as judged by Coomassie-stained SDS-PAGE gels of the eluted samples. Trace levels of ShhN binding to MmPtcT1 beads were detectable only by Western blotting; this binding was eliminated by EGTA. ShhN

and BOCFn13 visibly precipitate in the presence of LMWH and  $\text{Ca}^{2+}$ , resulting in high nonspecific binding to beads; we could therefore not evaluate specific binding of this complex to MmPtcT1. Since BOCFn13 and ShhN bind each other with micromolar affinity and each independently binds heparin, this precipitation was probably due to heparin-mediated protein crosslinking. Interactions between ShhN and MmPtcT1 beads were not enhanced by the presence of any BOC fragment, whether in the presence or absence of LMWH or  $\text{Ca}^{2+}$ . Binding of ShhN to MmPtcT1 beads was also unaffected by the presence of Hhip.

By Western blot analysis, ShhN did not adhere to MmPtcT1 resin to a greater degree than the negative control surface mutant ShhN-SC. However, ShhN-SC showed no dependence on EGTA and no precipitation in the presence of LMWH and HMT-BOCFn13, consistent with its inability to bind BOC (Figure 4.4).

HMT-BOCFn13 and HMT-BOCFn23 were not specifically pulled down by MmPtcT1-coupled beads. Although MmPtcT1-coupled beads did pull down BOC proteins as compared to blank beads (Figure 4.4), we observed that this binding was dependent on the presence of 1 mM  $\text{Ca}^{2+}$  and was eliminated by EGTA, despite the lack of known calcium binding to BOC proteins. We therefore performed additional control experiments comparing the binding and solution behaviors of tagged and untagged BOC proteins in the presence of calcium. These experiments showed that (i) His-tagged proteins oligomerize in the presence of 1 mM  $\text{Ca}^{2+}$  and that (ii) the binding of HMT-BOCFn13 to MmPtcT1 was dependent on the presence of His tags on both proteins. An irrelevant His-tagged protein was pulled down equally well by MmPtcT1 beads, and BOC proteins could also be pulled down equally by beads coupled with the irrelevant

His-tagged membrane protein AcrB. The presence of ShhN or ShhN-SC did not enhance the binding of BOC proteins to MmPtcT1 beads, nor did the addition of LMWH (Figure 4.4).

We also performed the pulldown experiments in reverse by coupling ShhN and ShhN-SC to beads and attempting to pull down detergent-solubilized MmPtcT1 alone or in the presence of HMT-BOCFn3 and HMT-BOCFn13. Hhip was added to test for its ability to block potential Ptc binding, and soluble ShhN and ShhN-SC were added as specific competitors for Ptc binding to ShhN- and ShhN-SC-coupled resins. Consistent with our previous results (Figure 4.4), we observed no binding above background of MmPtcT1 to ShhN resin. Aside from MmPtcT1, the resins pulled down the expected binding partners: ShhN resin pulled down BOCFn3 and HMT-BOCFn13 in a  $\text{Ca}^{2+}$  dependent manner, and pulldown of HMT-BOCFn3 by ShhN was blocked by Hhip.

*Pulldown of potential binding partners by DmPtcT1*

We next used similar pulldown assays to screen for DmPtcT1 binding to potential partners. DmPtcT1-coupled beads were used to pull down HhN, IhogFn12, IhogFn12 $\Delta$ H, IhogFn1 and IhogFn2. These proteins and their combinations were not found to interact strongly or specifically with DmPtcT1-coupled beads. Although IhogFn1, IhogFn12 and IhogFn12 $\Delta$ H were pulled down by DmPtcT1 beads in the absence of LMWH and could be detected in Coomassie-stained gels (Figure 4.5), this binding was only slightly greater than binding to the irrelevant membrane protein AcrB. In the presence of 200  $\mu\text{M}$  LMWH, mixing IhogFn1 or IhogFn12 with HhN induced visible precipitation, resulting in high non-specific binding to Strep-Tactin resin. We therefore could not evaluate specific binding to DmPtcT1 resin in these conditions. However, substituting 200  $\mu\text{M}$

decaheparin (DH) for LMWH prevented the precipitation of IhogFn1 and HhN and eliminated binding of IhogFn1 to DmPtcT1. HhN binding to DmPtcT1 beads was not detectable on Coomassie-stained gels except in the cases where precipitation was present. By Western blotting HhN did not bind DmPtcT1 beads specifically or to a greater degree than uncoupled Strep-Tactin beads. The presence of IhogFn1 and DH with HhN did not enhance the binding of HhN to DmPtcT1 beads.

Finally, we coupled HhN to amylose resin, and attempted to pull down soluble DmPtcT1. We also tested for binding to HhN of combinations of DmPtcT1 with IhogFn1, IhogFn2, IhogFn12, and IhogFn12 $\Delta$ H, and with soluble non-conjugated HhN (as a specific competitor for DmPtcT1 binding to HhN beads). No binding of DmPtcT1 was detectable in any condition, even when assessed by Western blotting. Aside from DmPtcT1, the HhN resin pulled down its known binding partners as expected: IhogFn1 and IhogFn12 in a heparin-dependent manner, and IhogFn12 $\Delta$ H much more weakly than IhogFn12. As expected, the HhN resin did not pull down IhogFn2.

#### *Co-expression and purification of Ptc proteins with potential binding partners*

Since *Drosophila* Ptc is believed to exist as a co-receptor complex with Ihog (Zheng, Mann, Sever, & Beachy, 2010), we investigated whether the co-expression and purification of IhogFn12<sup>TM</sup> with Ptc resulted in improved binding or better monodispersity of Ptc as judged by SEC. In addition, since the Hh proteins used in our pulldown assays were expressed in *E. coli*, they lacked lipid modifications. We therefore also investigated the co-expression and purification of Ptc proteins with dually lipidated Hh proteins, referred to as HhN<sub>p</sub> and ShhN<sub>p</sub>. These were expressed using dual-expression HhFL/Ski and ShhFL/Hhat baculoviruses, which simultaneously drive expression of the



full-length Hh proteins together with their respective palmitoyltransferases, although we did not determine whether HhN<sub>p</sub> or ShhN<sub>p</sub> were palmitoylated.

Co-expression of proteins did not result in co-purified stoichiometric complexes that interact strongly. When Sf9 cells were infected with a dual DmPtcT1/IhogFn12TM virus, both proteins were well-expressed. DmPtcT1 was purified using the standard Ni<sup>2+</sup>/Strep procedure (Figure 4.5A) and the resulting purified protein was separated by analytical SEC (Figure 4.5B). As judged by Western, some IhogFn12TM bound and eluted with DmPtcT1 during each purification step, but with significant amounts dissociating during each wash step. Comparison of the Ihog Western band intensities to standards (data not shown) indicated that the amount of IhogFn12TM remaining after DmPtcT1 purification was sub-stoichiometric with respect to DmPtcT1. The SEC behavior of DmPtcT1 co-expressed with IhogFn12TM was similar to that of DmPtcT1 expressed alone. Coomassie-stained SDS-PAGE and Western blotting of the SEC fractions indicated that DmPtcT1 and IhogFn12TM elution peaks overlapped, but with IhogFn12TM eluting at slightly higher apparent molecular weight.

When DmPtcT1 and HhN<sub>p</sub> were co-expressed, HhN<sub>p</sub> was observed to co-purify with DmPtcT1 (Figure 4.6A), with some losses during each affinity step. SEC followed by SDS-PAGE and Western blotting showed that the elution profile of HhN<sub>p</sub> was very broad, and its peak did not track the elution of DmPtcT1. Co-expression of all three proteins (DmPtcT1/IhogFn12TM with HhN<sub>p</sub>) did not appreciably alter the co-purification or SEC behavior of any protein.

Similarly, MmPtcT1 was co-expressed with ShhN<sub>p</sub> and purified (Figure 4.7A). Significant amounts of ShhN<sub>p</sub> remained in the flow-through during each affinity step, and

although some residual protein co-purified with MmPtcT1 and was detectable by Western, it was not visible on Coomassie-stained gels. By Western, the concentration of ShhN<sub>p</sub> clearly decreased during each affinity step, while the concentration of MmPtcT1 increased. The behavior of MmPtcT1 on SEC was not significantly altered by co-expression with ShhN<sub>p</sub>, and the elution peak of ShhN<sub>p</sub> did not match the peak of MmPtcT1 (Figure 4.7B).

## Discussion

In this study, mouse Patched-1 and *Drosophila* Patched proteins were expressed in Sf9 cells, detergent solubilized, purified, and characterized by SEC and CD. Deglycosylation and confocal microscopy studies suggested that Ptc proteins were glycosylated and transported to the cell surface, and Sonic Hedgehog from conditioned medium bound to MmPtcT1 in membranes isolated from MmPtcT1-expressing Sf9 cells. Detergent-extracted and purified Ptc proteins contained the expected secondary structure content as judged by CD but eluted as multiple species SEC with one peak consistent with a trimer or small oligomer and another of higher apparent molecular weight.

Despite our observations that Ptc proteins behaved as expected by CD, cell surface expression, and deglycosylation analysis, and that some purified protein ran at sizes potentially consistent with physiological oligomers by SEC, high-affinity binding to cognate Hh proteins or Ihog/BOC co-receptors could not be detected. Recent studies indicate that Ptc proteins have an absolute requirement for co-receptors to signal (BOC/CDO/Gas1 in mammals, Ihog/BOI in *Drosophila*), and that Ptc forms physical complexes with these co-receptors (Zheng, 2010; Izzi, 2011; Allen, 2011). In addition,

high-affinity binding of Shh to cells is Ptc-dependent (Fuse, 1999; Marigo, 1996). We were therefore surprised not to detect high-affinity binding between purified Ptc proteins and any of these putative binding partners, either pairwise or in ternary combinations.

While some bound ShhN was detected in Western blots after saturating amounts were incubated with MmPtcT1 beads, the amount of ShhN detected was not convincingly above levels of nonspecific binding to blank beads (Figure 4.4). Based on the clearly observed binding of BOC/Ihog proteins to their cognate Hh proteins in our pulldown assays and their known  $K_d$ 's of  $\sim 2.5\text{--}5\ \mu\text{M}$ , we expect interactions between Ptc proteins and any partner would have resulted in the clear observation of Hh on Coomassie-stained gels if their  $K_d$  were below  $\sim 5\ \mu\text{M}$ . Furthermore, the surface mutant ShhN-SC, which has impaired binding to Ptc-expressing cells (Fuse, 1999), showed roughly the same trace levels of binding to MmPtcT1 beads as native ShhN. The addition of Hhip also did not compete with Ptc beads for ShhN binding, further suggesting a lack of specific binding of ShhN to MmPtcT1 beads. Thus, we were not able to reproduce a previous observation of high-affinity binding between ShhN and purified mouse Ptc (Joubert, 2009). In addition, HMT-BOCFn13 and HMT-BOCFn23 did not bind specifically to MmPtcT1. Binding in initial studies in the presence of  $\text{Ca}^{2+}$  proved to be mediated by His tags (all BOC and Ptc proteins, but not Ihog proteins, were His-tag fusions) in the presence of  $\text{Ca}^{2+}$  ions. The presence of Shh together with BOC fragments did not alter the apparent binding of any of these proteins to MmPtcT1, providing no evidence for the presence of a ternary complex between soluble forms of BOC and Shh and detergent-solubilized Ptc.

*Drosophila* HhN also did not bind DmPtcT1 beads at levels appreciably above background, as judged by Western blotting. No HhN was detectable on Coomassie-

stained gels. IhogFn1, IhogFn12, and IhogFn12 $\Delta$ H were bound by DmPtcT1 beads in the absence of both heparin and Hh, as detected by Coomassie-stained gels; however, for IhogFn12, roughly comparable binding was also seen to beads coupled with the irrelevant membrane protein AcrB. The addition of HhN and heparin decasachharide eliminated binding of IhogFn1 to DmPtcT1 beads. In contrast with previous observations (Zheng, 2010), no binding of IhogFn2 to DmPtcT1 beads was observed. In all cases where some interaction was observed, the relative stoichiometry of the bound protein to Ptc was low, as judged by comparing Coomassie band intensities. Given the low stoichiometry of all observed complexes and the nonspecific binding seen between IhogFn12 and AcrB, we conclude that specific binding of Ihog fragments to Ptc was not observed.

As a final effort at purifying high-affinity, stoichiometric complexes of Ptc proteins with putative ligands, we attempted several co-expression experiments: MmPtcT1 with ShhN<sub>p</sub>, DmPtcT1 with HhN<sub>p</sub>, DmPtcT1 with IhogFn12TM, and DmPtcT1 with HhN<sub>p</sub> and IhogFn12TM. In all cases, as Ptc proteins were purified using His and SBP tags, their expression partners increasingly dissociated with each chromatographic step. A final Superose6 chromatography step showed some proportion of binding partner co-migrating with Ptc in each case, but with most having been lost in preceding purification steps. Qualitatively, this suggests at most a weak interaction between purified Ptc and these binding partners.

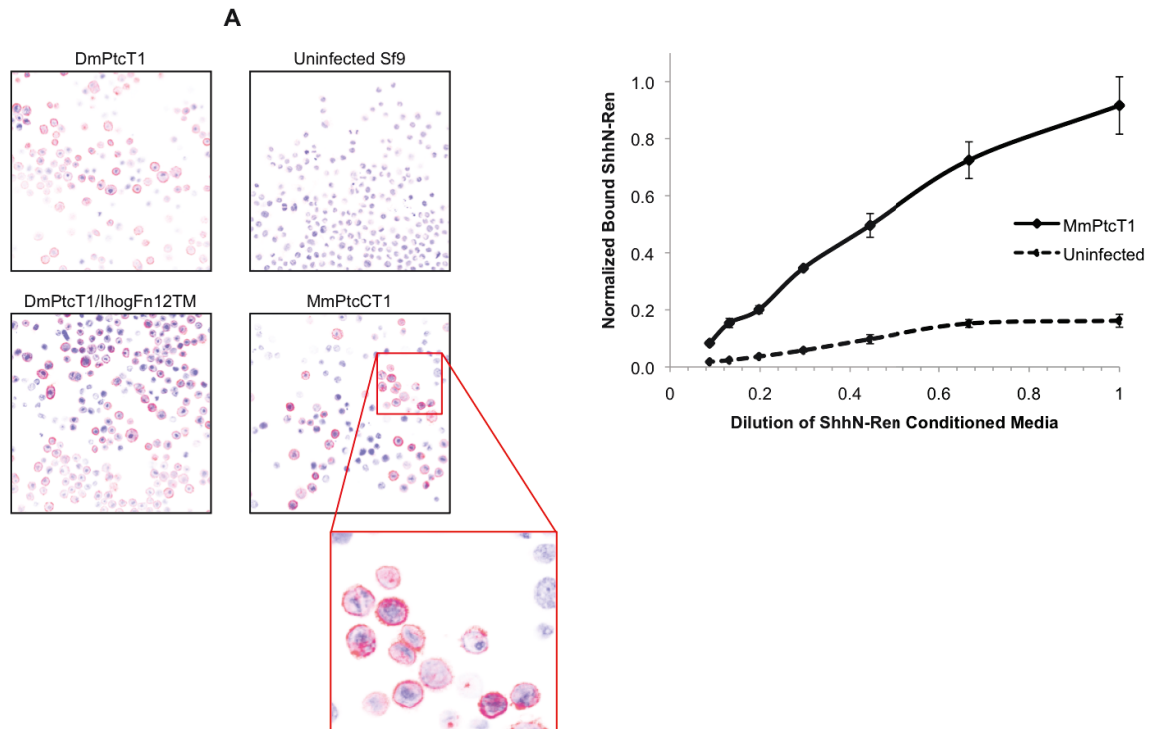
The absence of high-affinity interactions between purified Ptc and several potential binding partners observed here could be explained in several ways. Firstly, Ptc may not adopt a native or binding-competent conformation in this expression system, or it may lose these properties during detergent solubilization. The ability of both Ptc proteins to

reach the cell surface and of MmPtcT1 to become glycosylated and bind ShhN when Sf9 cell membranes suggests initial folding and biogenesis occurs properly (Figure 4.1), however, and that any non-native structure is more likely to stem from detergent solubilization. Regrettably, no functional assay exists to validate the integrity of purified Ptc—indeed, the central objective of our studies was to find an *in vitro* activity for Ptc that could be studied—and we can only rely our CD and SEC results to conclude that purified Ptc retains at least some native-like structure.

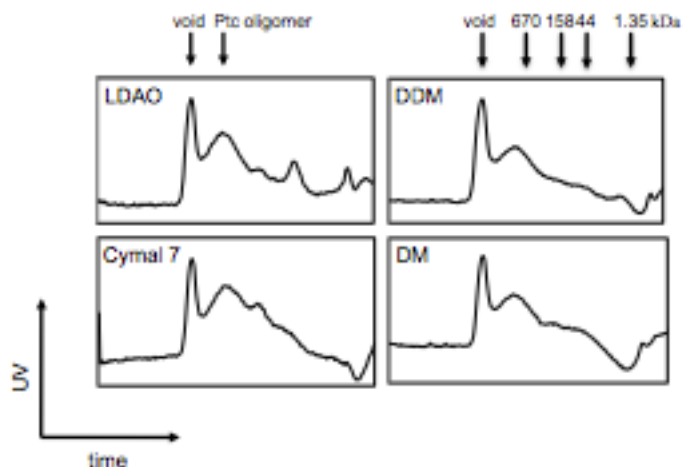
Although fos-choline detergents are generally considered harsher detergents that may destabilize membrane proteins during solubilization—a particular concern for proteins like Ptc that are only efficiently extracted by fos-choline detergents—many membrane proteins retain function in fos-choline. Proteins have been crystallized directly from fos-choline detergents, for example, including alpha helical membrane proteins [43]. The transmembrane protein SCAP was also only extracted well by a fos-choline detergent and was shown to retain function in detergent-extracted form. SCAP is similar to Ptc in being a multi-pass transmembrane protein (a tetramer of 8-pass subunits) with a sterol-sensing domain and was also expressed in Sf9 cells with N-terminal tags. If Hh proteins interact with a region of the large, presumably independently-folded extracellular loops of Ptc as has been suggested, these regions of Ptc may also be less susceptible to destabilization by detergents than the Ptc transmembrane regions.

Another possibility for the absence of strong binding between Ptc and cognate Hh and CDO/Ihog proteins is that additional cell-surface co-factors not present in our assays are required for high-affinity interactions or to maintain Ptc stability. The low concentration of luciferase-tagged Shh in the conditioned medium used to show binding between Shh

and mPtc in isolated Sf9 membranes precluded saturation of binding and estimation of a binding constant, but the likely absence of mammalian co-factors in these cells suggests that mPtc itself can bind Shh with measurable affinity in the absence of known co-factors. A requirement for a bilayer environment, particular lipids, or glycans remains a possibility, however. Our inability to observe high-affinity binding between Ptc and cognate Hh or Ihog/CDO homologs thus seems likely to stem from loss of native-like Ptc function during detergent solubilization or detergent disruption of ShhN binding ability. As biochemical characterization of the nature and components of the functional Hh receptor remains a key objective of studies aiming to understand Hh signaling in normal and disease states, it will be important aspect of future efforts will be to find purification conditions that preserve or restore Ptc function. Nonetheless, the ability to produce near milligram amounts of purified Ptc described here establishes a cornerstone for future efforts to characterize the Hh receptor complex and elucidate the role of Ptc in Hh signaling

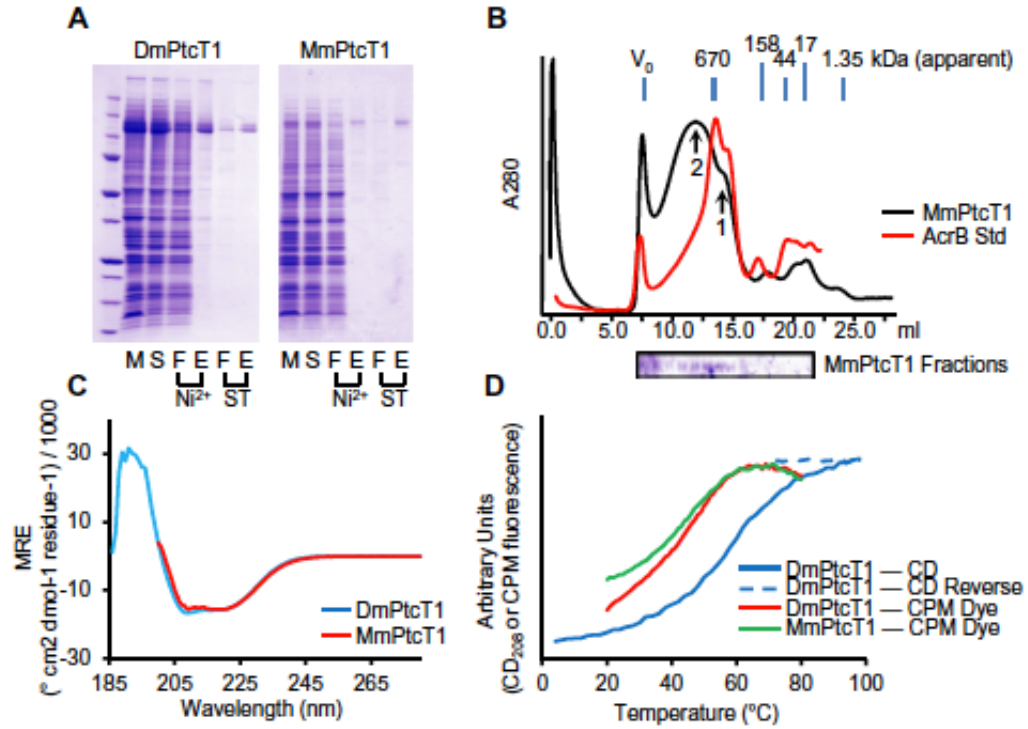


**Figure 4.1 Functional Ptc expresses on the surface of Sf9 cells.** (A) Sf9 cells infected with the indicated viruses were immunostained using anti-Myc (9E10) primary and Cy3 secondary antibodies (red), and imaged by confocal microscopy. Nuclei were stained with DAPI (blue). (B) Membranes from uninfected Sf9 cells or cells infected with MmPtcT1 virus were isolated, incubated with varying concentrations of ShhN-Renilla luciferase conditioned medium, washed, and residual luciferase activities determined.

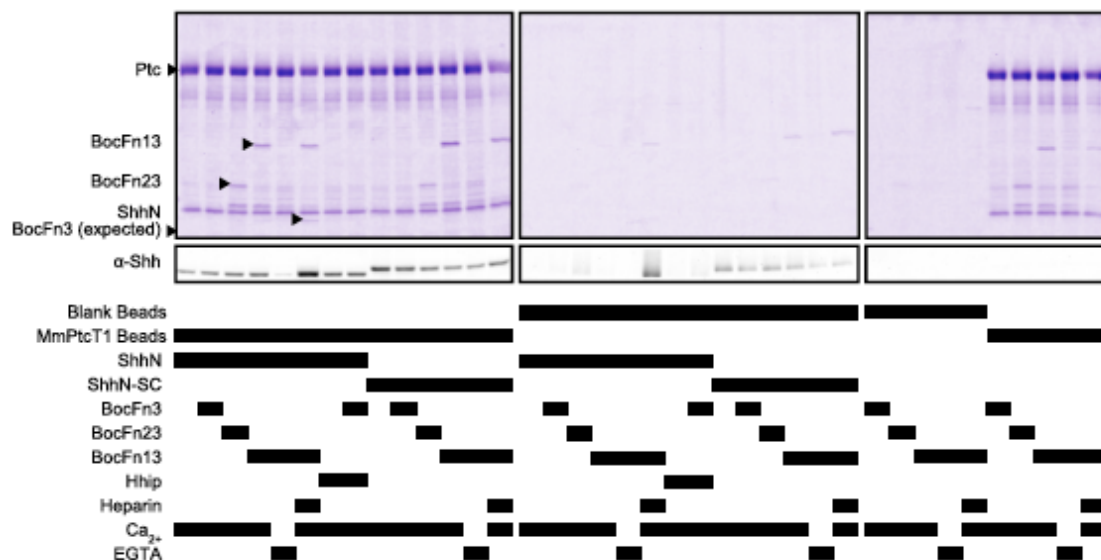


**Figure 4.2 SEC of Ptc proteins after exchange into different detergents.** Superose6 size-exclusion chromatograms of MmPtcT1 solubilized in fos-choline 12 and exchanged during purification into the indicated detergents are shown. Similar behavior after initial purification was observed in most detergents and for DmPtcT1 (not shown), with an oligomeric Ptc peak eluting slightly before the 670 kDa marker, and in some cases another peak between the 158 and 670 kDa markers

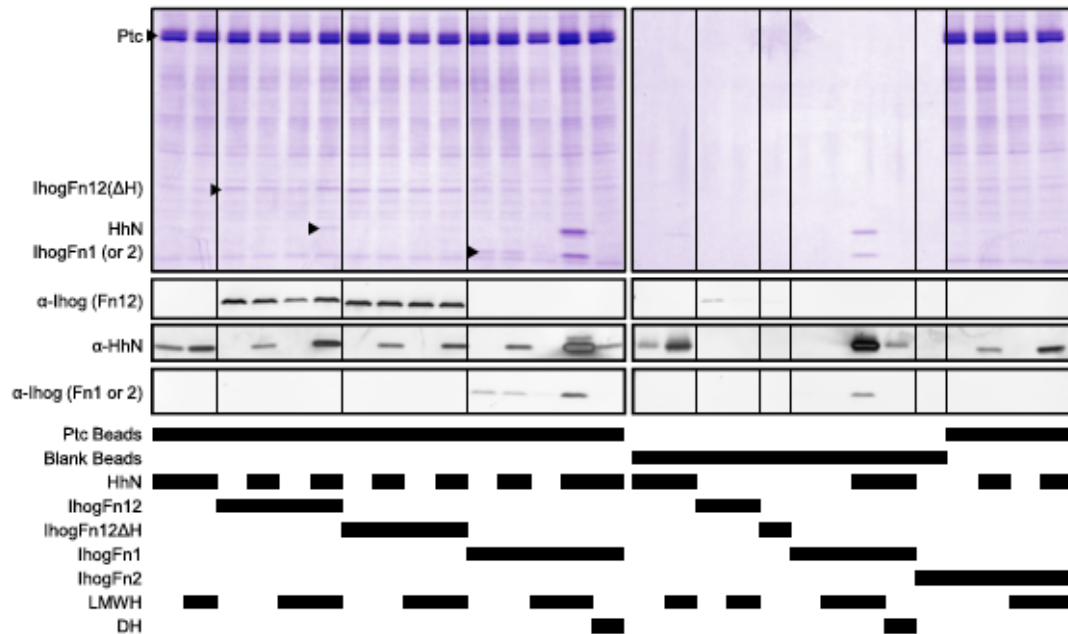




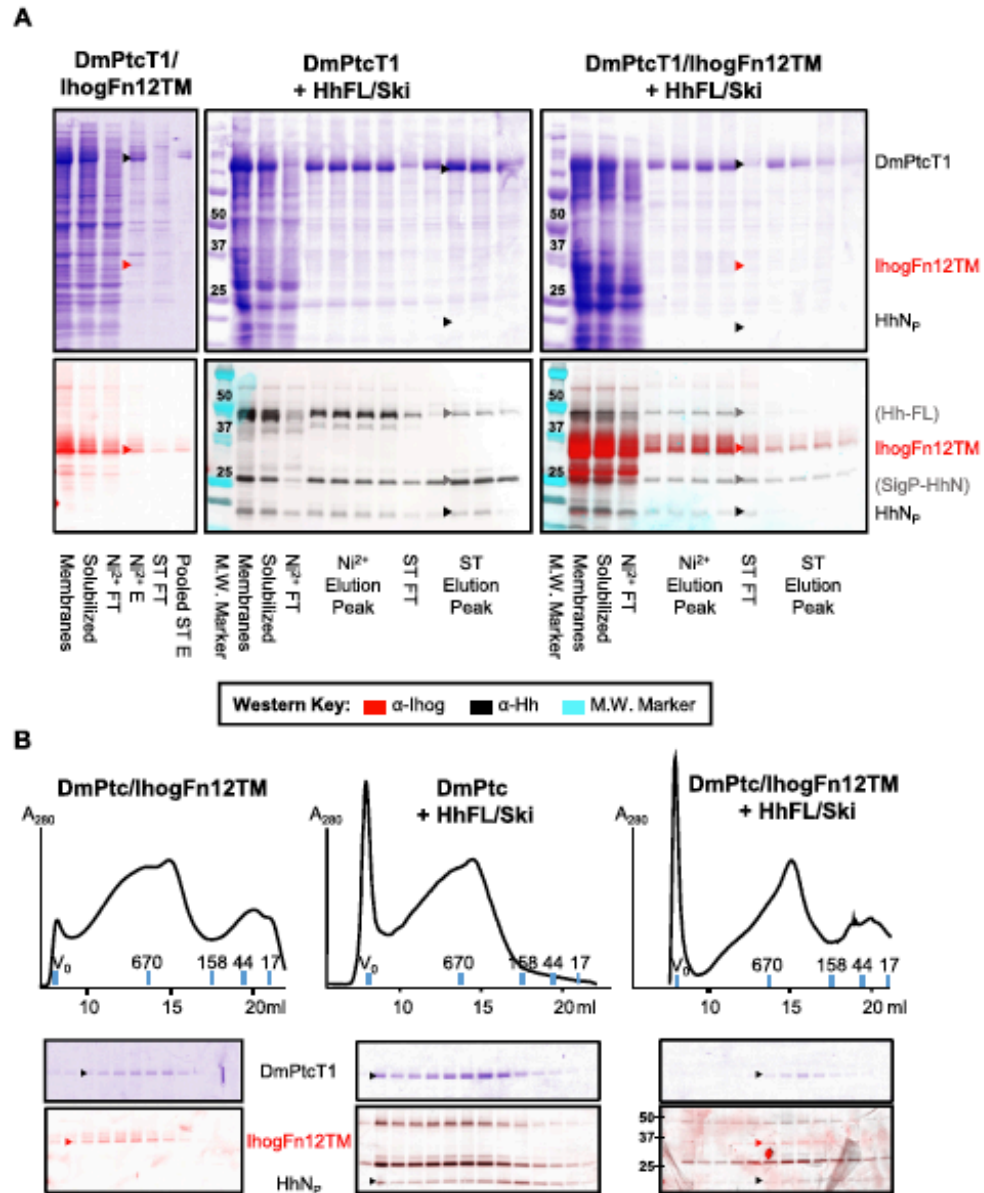
**Figure 4.3 Purification and characterization of Ptc proteins.** (A) DmPtcT1 and MmPtcT1 were purified by successive  $\text{Ni}^{2+}$  and Strep-Tactin (ST) chromatography steps. Gel Lanes: (M) resuspended membranes; (S) detergent-solubilized membranes; (F) flow-through; (E) elution (B) SEC of purified, concentrated MmPtcT1 (140 kDa calculated) compared to tagged AcrB (123 kDa), a protein standard with the same predicted topology. MmPtcT1 elutes in two peaks: a small oligomer consistent with a potential physiological trimer (arrow 1) and a larger oligomer (arrow 2) (C) CD spectra of DmPtcT1 and MmPtcT1 show substantial alpha-helical character. (D) Thermal unfolding of Ptc proteins as measured by CD at 208 nm (correlated with alpha helical content) and by CPM dye binding to exposed cysteines.



**Figure 4.4 Pulldown of potential binding partners by MmPtcT1.** The indicated proteins and cofactors were incubated with MmPtcT1-coupled Strep-Tactin resin and blank (uncoupled) Strep-Tactin resin. Bound proteins were eluted with biotin and analyzed by Coomassie-stained SDS-PAGE.

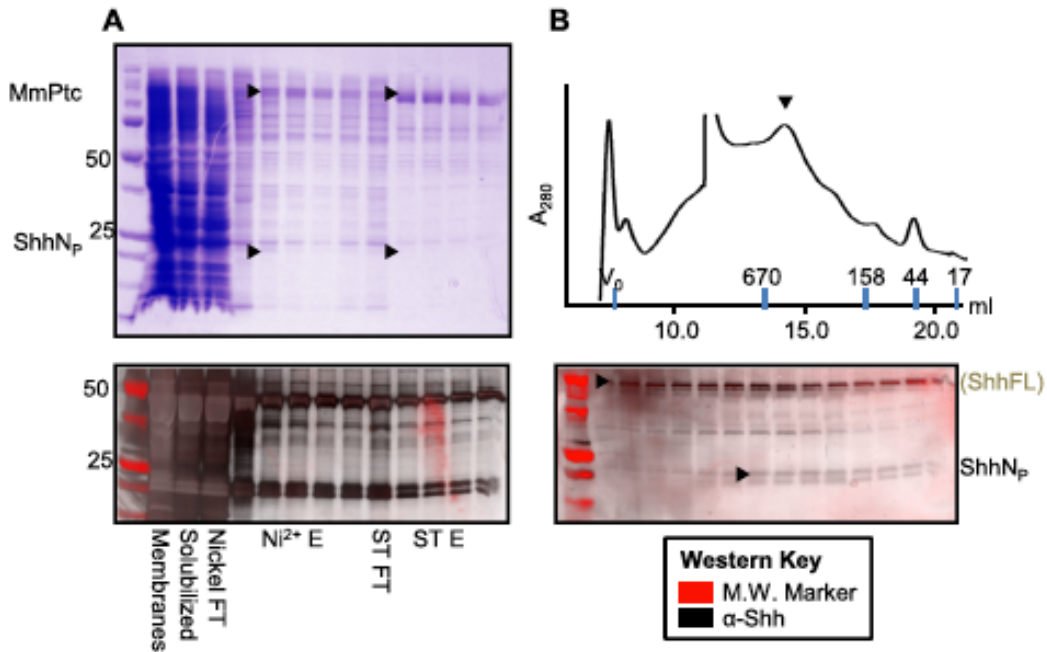


**Figure 4.5 Pulldown of potential binding partners by DmPtcT1.** Proteins and other cofactors were incubated with DmPtcT1-coupled Strep-Tactin resin and blank (uncoupled) Strep-Tactin resin in the indicated combinations. Bound proteins were eluted with biotin and analyzed by Coomassie-stained SDS-PAGE and anti-Shh Western blot. Shh and BOC fragments, marked with arrows, that appear to co-precipitate with Ptc proved to be due to heparin-dependent precipitation (Shh) or His-tag and  $\text{Ca}^{2+}$  dependent association (BOC).



**Figure 4.6 Co-expression and purification of DmPtcT1 with potential binding partners.** Sf9 cells were co-infected with baculoviruses driving expression of DmPtcT1 along with IhogFn12TM, HhFL/Ski, or both IhogFn12TM and HhFL/Ski. Expected positions of indicated protein bands are marked. (A) DmPtcT1 was purified and detected by Coomassie-stained SDS-PAGE. The presence of co-purifying Hh or Ihog was followed by Western blot. The same Western blot membrane was sequentially probed

with Ihog and Hh antibodies. The images were then overlaid and color-coded (see key). Three major Hh bands are observed: one corresponding to the molecular weight of unprocessed Hh (Hh-FL), one corresponding to C-terminally processed Hh retaining its signal sequence (SigP-HhN), and one corresponding to HhNp. (B) The resulting co-purified proteins were subjected to Superose6 SEC, and the fractions analyzed by Coomassie-stained SDS-PAGE (to visualize DmPtcT1) and by Western blot, probing sequentially for Hh and Ihog as in (A).



**Figure 4.7 Co-expression and purification of MmPtcT1 with Shh.** Sf9 cells were co-infected with baculoviruses driving expression of MmPtcT1 along with ShhFL/Hhat. (A) MmPtcT1 was purified and visualized with Coomassie-stained SDS-PAGE. The presence of co-purifying Shh followed by Western blot. Two major Shh bands are observed: one corresponding to the molecular weight of the unprocessed Shh (ShhFL), and one corresponding to the expected molecular weight of ShhN<sub>p</sub>. (B) The resulting co-purified proteins were subjected to Superose6 SEC. The fractions were analyzed by Coomassie-stained SDS-PAGE for MmPtcT1 detection, and by Western blot for Shh.

## Bibliography

- Accili, D., Drago, J., Lee, E. J., Johnson, M. D., Cool, M. H., Salvatore, P., Asico, L. D., Jose, P. A., Taylor, S. I., and Westphal, H. (1996) Early neonatal death in mice homozygous for a null allele of the insulin receptor gene. *Nat Genet* **12**, 106-109.
- Aeed, P. A., & Elhammer, A. P. (1994). Glycosylation of recombinant prorenin in insect cells: the insect cell line Sf9 does not express the mannose 6-phosphate recognition signal. *Biochemistry* **33**, 8793-7.
- Alcedo, J., Ayzenzon, M., Ohlen, Von, T., Noll, M., and Hooper, J. E. (1996) The *Drosophila* smoothened gene encodes a seven-pass membrane protein, a putative receptor for the hedgehog signal. *Cell* **86**, 221-232.
- Alexandrov, A. I., Mileni, M., Chien, E. Y., Hanson, M. A., & Stevens, R. C. (2008). Microscale fluorescent thermal stability assay for membrane proteins. *Structure* **16**, 351-9.
- Allen, B. L., Song, J. Y., Izzi, L., Althaus, I. W., Kang, J.-S., Charron, F., Krauss, R. S., and McMahon, A. P. (2011) Overlapping roles and collective requirement for the coreceptors GAS1, CDO, and BOC in SHH pathway function. *Dev. Cell* **20**, 775-787.
- Alvarado, D., Klein, D. E. and Lemmon, M. A. (2009) ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. *Nature* **461**, 287-291.
- Alvarado, D., Klein, D. E., and Lemmon, M. A. (2010) Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell* **142**, 568-579.
- Arkhipov, A., Shan, Y., Das, R., Endres, N. F., Eastwood, M. P., Wemmer, D. E., Kuriyan, J., and Shaw, D. E. (2013) Architecture and membrane interactions of the EGF receptor. *Cell* **152**, 557-569.
- Arwert, E. N., Hoste, E., and Watt, F. M. (2012) Epithelial stem cells, wound healing and cancer. *Nat. Rev. Cancer* **12**, 170-180.
- Ayer, K. L., and Therond, P. P. (2010) Evaluating Smoothened as a G-protein-coupled receptor for Hedgehog signalling. *Trends Cell Biol.* **20**, 287-298.
- Baer, K., Al-Hasani, H., Parvaresh, S., Corona, T., Rufer, A., Nolle, V., Bergschneider, E., and Klein, H.W. (2001) Dimerization-induced activation of soluble insulin/IGF-1 receptor kinases: an alternative mechanism of activation. *Biochemistry* **40**, 14268-14278.
- Bahadur, R.P., Chakrabarti, P., Rodier, F., and Janin, J. (2004) A dissection of specific and non-specific protein-protein interfaces. *J Mol Biol* **336**, 943-955.

- Ballesteros, J. A., and Weinstein, H. (1995) in *Methods in Neurosciences* (Sealfon, S. C., ed.) pp. 366–428, Academic Press.
- Barak, L. S., Ménard, L., Ferguson, S. S., Colapietro, A. M., and Caron, M. G. (1995) The conserved seven-transmembrane sequence NP(X)<sub>2</sub>Y of the G-protein-coupled receptor superfamily regulates multiple properties of the beta 2-adrenergic receptor. *Biochemistry* **34**, 15407–15414.
- Bass, J., Kurose, T., Pashmforoush, M., and Steiner, D.F. (1996) Fusion of insulin receptor ectodomains to immunoglobulin constant domains reproduces high-affinity insulin binding in vitro. *J Biol Chem* **271**, 19367–19375.
- Bazan, J. F., and de Sauvage, F. J. (2009) Structural Ties between Cholesterol Transport and Morphogen Signaling. *Cell* **138**, 1055–1056.
- Beachy, P. A., Hymowitz, S. G., Lazarus, R. A., Leahy, D. J., and Siebold, C. (2010) Interactions between Hedgehog proteins and their binding partners come into view. *Genes Dev.* **24**, 2001–2012.
- Beglev, D, Roux, B. 1994. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *The Journal of Chemical Physics* **100**:9050-9063.
- Bender, M. H., Hipskind, P. A., Capen, A. R., Cockman, M., Credille, K. M., Gao, H., Bastian, J. A., Clay, J. M., Lobb, K. L., Sall, D. J., Thompson, M. L., Wilson, T., Wishart, G. N., and Patel, B. K. R. (2011) Abstract 2819: Identification and characterization of a novel smoothened antagonist for the treatment of cancer with deregulated hedgehog signaling. *Cancer Res.* **71**, 2819–2819.
- Bijlsma, M. F., Spek, C. A., Zivkovic, D., van de Water, S., Rezaee, F., and Peppelenbosch, M. P. (2006) Repression of Smoothened by Patched-Dependent (Pro-)Vitamin D3 Secretion. *PLoS Biol.* **4**, 1397–1410.
- Bishop, B., Aricescu, A. R., Harlos, K., O'Callaghan, C. A., Jones, E. Y., and Siebold, C. (2009) Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP. *Nat. Struct. Mol. Biol.* **16**, 698–703.
- Bosanac, I., Maun, H. R., Scales, S. J., Wen, X., Lingel, A., Bazan, J. F., de Sauvage, F. J., Hymowitz, S. G., and Lazarus, R. A. (2009) The structure of SHH in complex with HHIP reveals a recognition role for the Shh pseudo active site in signaling. *Nat. Struct. Mol. Biol.* **16**, 691–697.
- Brastianos, P. K., Horowitz, P. M., Santagata, S., Jones, R. T., McKenna, A., Getz, G., Ligon, K. L., Palescandolo, E., Van Hummelen, P., Ducar, M. D., Raza, A., Sunkavalli, A., Macconail, L. E., Stemmer-Rachamimov, A. O., Louis, D. N., Hahn, W. C., Dunn, I.



F., and Beroukhi, R. (2013) Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. *Nat. Genet.* **45**, 285–289.

Brinkhuizen, T., Reinders, M. G., van Geel, M., Hendriksen, A. J. L., Paulussen, A. D. C., Winnepenninckx, V. J., Keymeulen, K. B., Soetekouw, P. M. M. B., van Steensel, M. A. M., and Mosterd, K. (2014) Acquired resistance to the Hedgehog pathway inhibitor vismodegib due to smoothened mutations in treatment of locally advanced basal cell carcinoma. *J. Am. Acad. Dermatol.*, doi: 10.1016-j.jaad.2014.08.001.

Briscoe, J., and Therond, P. P. (2013) The mechanisms of Hedgehog signalling and its roles in development and disease. *Nat. Rev. Mol. Cell Biol.* **14**, 418–431.

Buonamici, S., Williams, J., Morrissey, M., Wang, A., Guo, R., Vattay, A., Hsiao, K., Yuan, J., Green, J., Ospina, B., Yu, Q., Ostrom, L., Fordjour, P., Anderson, D. L., Monahan, J. E., Kelleher, J. F., Peukert, S., Pan, S., Wu, X., Maira, S.-M., García-Echeverría, C., Briggs, K. J., Watkins, D. N., Yao, Y.-M., Lengauer, C., Warmuth, M., Sellers, W. R., and Dorsch, M. (2010) Interfering with resistance to smoothened antagonists by inhibition of the PI3K pathway in medulloblastoma. *Sci Transl Med* **2**, 51ra70.

Burgess, A.W., Cho, H.S., Eigenbrot, C., Ferguson, K.M., Garrett, T.P., Leahy, D.J., Lemmon, M.A., Sliwkowski, M.X., Ward, C.W., and Yokoyama, S. (2003) An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol Cell* **12**, 541–552.

Cabail, M. Z., Li, S., Lemmon, E., Bowen, M. E., Hubbard, S. R., and Miller, W. T. (2015) The insulin and IGF1 receptor kinase domains are functional dimers in the activated state. *Nat Commun* **6**, 6406.

Cao, H., Bangalore, L., Dompe, C., Bormann, B.J., and Stern, D.F. (1992) An extra cysteine proximal to the transmembrane domain induces differential cross-linking of p185neu and p185neu. *J Biol Chem* **267**, 20489–20492.

Carbe, C. J., Cheng, L., Addya, S., Gold, J. I., Gao, E., Koch, W. J., and Riobo, N. A. (2014) Gi proteins mediate activation of the canonical hedgehog pathway in the myocardium. *AJP: Heart and Circulatory Physiology* **307**, H66–H72.

Carroll, C. E., Marada, S., Stewart, D. P., Ouyang, J. X., and Ogden, S. K. (2012) The extracellular loops of Smoothened play a regulatory role in control of Hedgehog pathway activation. *Development* **139**, 827–827.

Chang, A. L. S., and Oro, A. E. (2012) Initial assessment of tumor regrowth after vismodegib in advanced Basal cell carcinoma. *Arch Dermatol* **148**, 1324–1325.

- Chaudhary, S., Pak, J. E., Gruswitz, F., Sharma, V., and Stroud, R. M. (2012). Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. *Nat Protocols* **7**, 453-66.
- Cheatham B., Shoelsen, S. E., Yamada, K., and Kahn, C. R. (1993) Substitution of the erbB-2 oncoprotein transmembrane domain activates the insulin receptor and modulates the action of the insulin and insulin-receptor substrate 1. *Proceedings of the National Academy of Sciences of USA* **90**:7336-7340.
- Cheatham, B., and Kahn, C.R. (1992) Cysteine 647 in the insulin receptor is required for normal covalent interaction between alpha- and beta-subunits and signal transduction. *J Biol Chem* **267**, 7108-7115.
- Chen, J. K., Taipale, J., Young, K. E., Maiti, T., and Beachy, P. A. (2002) Small molecule modulation of Smoothened activity. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14071–14076
- Chen, L., Novicky, L., Merzlyakov, M., Hristov, T., and Hristova, K. (2010a) Measuring the energetics of membrane protein dimerization in mammalian membranes. *J Am Chem Soc* **132**, 3628-3635.
- Chen, L., Placone, J., Novicky, L., and Hristova, K. (2010b) The extracellular domain of fibroblast growth factor receptor 3 inhibits ligand-independent dimerization. *Sci Signal* **3**, ra86.
- Chen, M. H., Li, Y. J., Kawakami, T., Xu, S. M., and Chuang, P. T. (2004) Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. *Genes Dev* **18**, 641-59.
- Chen, W. (2004) Activity-Dependent Internalization of Smoothened Mediated by - Arrestin 2 and GRK2. *Science* **306**, 2257–2260.
- Chen, Y., Sasai, N., Ma, G., Yue, T., Jia, J., Briscoe, J., and Jiang, J. (2011) Sonic Hedgehog dependent phosphorylation by CK1 $\alpha$  and GRK2 is required for ciliary accumulation and activation of smoothened. *PLoS Biol.* **9**, e1001083.
- Chuang, P. T. and McMahon, A. P. (1999) Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* **397**, 617-621.
- Chung, I., Akita, R., Vandlen, R., Toomre, D., Schlessinger, J. and Mellman, I. (2010) Spatial control of EGF receptor activation by reversible dimerization in living cells. *Nature* **464**, 783-787.
- Clark, V. E., Erson-Omay, E. Z., Serin, A., Yin, J., Cotney, J., Ozduman, K., Avşar, T., Li, J., Murray, P. B., Henegariu, O., Yilmaz, S., Günel, J. M., Carrión-Grant, G., Yilmaz, B., Grady, C., Tanrikulu, B., Bakircioğlu, M., Kaymakçalan, H., Caglayan, A. O., Sencar,

L., Ceyhun, E., Atik, A. F., Bayri, Y., Bai, H., Kolb, L. E., Hebert, R. M., Omay, S. B., Mishra-Gorur, K., Choi, M., Overton, J. D., Holland, E. C., Mane, S., State, M. W., Bilgüvar, K., Baehring, J. M., Gutin, P. H., Piepmeier, J. M., Vortmeyer, A., Brennan, C. W., Pamir, M. N., Kiliç, T., Lifton, R. P., Noonan, J. P., Yasuno, K., and Günel, M. (2013) Genomic analysis of non-NF2 meningiomas reveals mutations in TRAF7, KLF4, AKT1, and SMO. *Science* **339**, 1077–1080.

Cooper, M. K., Porter, J. A., Young, K. E., and Beachy, P. A. (1998) Teratogen-Mediated Inhibition of Target Tissue Response to Shh Signaling. *Science* **280**, 1603–1607.

Cooper, M. K., Wassif, C. A., Krakowiak, P. A., Taipale, J., Gong, R., Kelley, R. I., Porter, F. D., and Beachy, P. A. (2003) A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nat. Genet.* **33**, 508–513.

Corcoran, R. B., and Scott, M. P. (2006) Oxysterols stimulate Sonic hedgehog signal transduction and proliferation of medulloblastoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8408–8413.

Craddock, B.P., Cotter, C., and Miller, W.T. (2007) Autoinhibition of the insulin-like growth factor I receptor by the juxtamembrane region. *FEBS Lett* **581**, 3235–3240.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190.

De Meyts, P. (2004) Insulin and its receptor: structure, function and evolution. *Bioessays* **26**, 1351–1362.

De Meyts, P., and Whittaker, J. (2002) Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat Rev Drug Discov* **1**, 769–783.

Del Piccolo, N., Placone, J., He, L., Agudelo, S.C., and Hristova, K. (2012) Production of plasma membrane vesicles with chloride salts and their utility as a cell membrane mimetic for biophysical characterization of membrane protein interactions. *Anal Chem* **84**, 8650–8655.

Denef, N., Neubuser, D., Perez, L., and Cohen, S. M. (2000) Hedgehog Induces Opposite Changes in Turnover and Subcellular Localization of Patched and Smoothed. *Cell* **102**, 521–531.

Depetris R. S., Hu J., Gimeplevich I., Holt L. J., Daly R. J., and Hubbard, S. R. (2005) Structural basis for inhibition of the insulin receptor by the adaptor protein Grb14. *Mol Cell* **20**, 325–333.

Deupi, X., and Standfuss, J. (2011) Structural insights into agonist-induced activation of G-protein-coupled receptors. *Curr. Opin. Struct. Biol.* **21**, 541–551.

- Deupi, X., Edwards, P., Singhal, A., Nickle, B., Oprian, D., Schertler, G., and Standfuss, J. (2012) Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 119–124.
- Dwyer, J. R., Sever, N., Carlson, M., Nelson, S. F., Beachy, P. A., and Parhami, F. (2007) Oxysterols Are Novel Activators of the Hedgehog Signaling Pathway in Pluripotent Mesenchymal Cells. *J. Biol. Chem.* **282**, 8959–8968.
- Endres, N.F., Das, R., Smith, A.W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J.G., Shan, Y., Shaw, D.E., Wemmer, D.E., et al. (2013) Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543-556.
- Favelyukis, S., Till, J.H., Hubbard, S.R., and Miller, W.T. (2001). Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol* **8**, 1058-1063.
- Forbes, A. J., Nakano, Y., Taylor, A. M., and Ingham, P. W. (1993) Genetic analysis of hedgehog signalling in the Drosophila embryo. *Development Supplement*, 115-24.
- Fratalli A. L., Treadyway J. L., and Pessin J. E. (1991) Evidence supporting a passive role for the insulin receptor transmembrane domain in insulin-dependent signal transduction. *J Biol Chem* **267**, 19521-19528.
- Frattali, A. L., Treadway, J. L., and Pessin, J. E. (1992) Transmembrane signaling by the human insulin receptor kinase. Relationship between intramolecular beta subunit trans- and cis-autophosphorylation and substrate kinase activation. *J Biol Chem* **267**, 19521-19528.
- Fuse, N., Maiti, T., Wang, B., Porter, J. A., Hall, T. M. T., Leahy, D. J., and Beachy, P. A. (1999) Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for Patched. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10992–10999.
- Geisbrecht, B. V., Bouyain, S., & Pop, M. (2006). An optimized system for expression and purification of secreted bacterial proteins. *Protein Expression and Purification* **46**, 23-32.
- Goodrich, L. V., Milenkovic, L., Higgins, K. M., and Scott, M. P. (1997) Altered Neural Cell Fates and Medulloblastoma in Mouse patched Mutants. *Science* **277**, 1109–1113.
- Gradilla, A.-C., and Guerrero, I. (2013) Hedgehog on the move: a precise spatial control of Hedgehog dispersion shapes the gradient. *Curr. Opin. Genet. Dev.* **23**, 363–373.
- Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V.-P., Chien, E. Y. T., Velasquez, J., Kuhn, P., and Stevens, R. C. (2008) A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* **16**, 897–905.

Hanson, M. A., Roth, C. B., Jo, E., Griffith, M. T., Scott, F. L., Reinhart, G., Desale, H., Clemons, B., Cahalan, S. M., Schuerer, S. C., Sanna, M. G., Han, G. W., Kuhn, P., Hugh Rosen, and Stevens, R. C. (2012) Crystal Structure of a Lipid G Protein–Coupled Receptor. *Science* **335**, 851–855.

Heemskerk, J., and DiNardo, S. (1994) Drosophila hedgehog acts as a morphogen in cellular patterning. *Cell* **76**, 449–460.

Hidalgo, A., and Ingham, P. (1990) Cell patterning in the Drosophila segment: spatial regulation of the segment polarity gene patched. *Development* **110**, 291–301.

Hoover, W.G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Physical Review A* **31**, 1695–1697.

Houde, D., and Demarest, S.J. (2011) Fine details of IGF-1R activation, inhibition, and asymmetry determined by associated hydrogen /deuterium-exchange and peptide mass mapping. *Structure* **19**, 890–900.

Hoyne, P. A., Cosgrove, L. J., McKern, N. M., Bentley, J. D., Ivancic, N., Elleman, T. C., and Ward, C. W. (2000) High affinity insulin binding by soluble insulin receptor extracellular domain fused to a leucine zipper. *FEBS Lett* **479**, 15–18.

Hu, J., Liu, J., Ghirlando, R., Saltiel, A. R., Hubbard, S. R. (2003) Structural basis for recruitment of the adaptor protein APS to the activated insulin receptor. *Mol Cell* **12**, 1379–1389.

Huangfu, D., Liu, A., Rakeman, A. S., Murcia, N. S., Niswander, L., and Anderson, K. V. (2003) Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* **426**, 83–87.

Hubbard, S.R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J* **16**, 5572–5581.

Hubbard, S.R. (2013) Structural biology: Insulin meets its receptor. *Nature* **493**, 171–172.

Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754.

Hui, C.-C., and Angers, S. (2011) Gli Proteins in Development and Disease. *Annu. Rev. Cell Dev. Biol.* **27**, 513–537.

Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell* **109**, 275–282.

Ingham, P. W. (2012). Hedgehog signaling. *CSH Pers Biol* **4**, a011221.

- Ingham, P. W., Nakano, Y., and Seger, C. (2011) Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat Rev Genet* **12**, 393–406.
- Ingham, P. W., Taylor, A. M., and Nakano, Y. (1991) Role of the Drosophila patched gene in positional signalling. *Nature* **353**, 184–187.
- Inlow, D., Shauger, A., and Maiorella, B. (1989) Insect cell culture and baculovirus propagation in protein-free medium. *Journal of Tissue Culture Methods* **12**, 13–16.
- Izzi, L., Lévesque, M., Morin, S., Laniel, D., Wilkes, B. C., Mille, F., Krauss, R. S., McMahon, A. P., Allen, B. L., and Charron, F. (2011) Boc and Gas1 each form distinct Shh receptor complexes with Ptch1 and are required for Shh-mediated cell proliferation. *Dev. Cell* **20**, 788–801.
- Janda, C. Y., Waghray, D., Levin, A. M., Thomas, C., and Garcia, K. C. (2012) Structural Basis of Wnt Recognition by Frizzled. *Science* **337**, 59–64.
- Jia, J., Tong, C., Wang, B., Luo, L., and Jiang, J. (2004) Hedgehog signalling activity of Smoothened requires phosphorylation by protein kinase A and casein kinase I. *Nature* **432**, 1045–1050.
- Johnson, R. L., Rothman, A. L., Xie, J., Goodrich, L. V., Bare, J. W., Bonifas, J. M., et al. (1996) Human Homolog of patched, a Candidate Gene for the Basal Cell Nevus Syndrome. *Science* **272**, 1668–71.
- Joubert, O., Nehmé, R., Fleury, D., De Rivoyre, M., Bidet, M., and Polidori, A. (2009) Functional studies of membrane-bound and purified human Hedgehog receptor Patched expressed in yeast. *Biochim Biophys Acta* **1788**, 1813–21.
- Kasuga, M., Fujita-Yamaguchi, Y., Blithe, D.L., White, M.F., and Kahn, C.R. (1983) Characterization of the insulin receptor kinase purified from human placental membranes. *J Biol Chem* **258**, 10973–10980.
- Kavran, J. M., Ward, M. D., Oladosu, O. O., Mulepati, S., and Leahy, D. J. (2010). All mammalian Hedgehog proteins interact with cell adhesion molecule, down-regulated by oncogenes (CDO) and brother of CDO (BOC) in a conserved manner. *J Biol Chem* **285**, 24584–90.
- Keefe, A. D., Wilson, D. S., Seelig, B., and Szostak, J. W. (2001) One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expression and Purification* **23**, 440–6.
- Kim, J., Aftab, B. T., Tang, J. Y., Kim, D., Lee, A. H., Rezaee, M., Kim, J., Chen, B., King, E. M., Borodovsky, A., Riggins, G. J., Epstein, E. H., Jr., Beachy, P. A., and Rudin, C. M. (2013) Itraconazole and Arsenic Trioxide Inhibit Hedgehog Pathway

- Activation and Tumor Growth Associated with Acquired Resistance to Smoothed Antagonists. *Cancer Cell* **23**, 23–34.
- Kim, M.-S., Saunders, A. M., Hamaoka, B. Y., Beachy, P. A., and Leahy, D. J. (2011) Structure of the protein core of the glypican Dally-like and localization of a region important for hedgehog signaling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13112–13117.
- King, C., Sarabipour, S., Byrne, P.O., Leahy, D.J., and Hristova, K. (2014) The FRET signatures of noninteracting proteins in membranes: simulations and experiments. *Biophys J* **106**, 1-9.
- Kleiman, L.B., Maiwald, T., Conzelmann, H., Lauffenburger, D.A., and Sorger, P.K. (2011) Rapid phospho-turnover by receptor tyrosine kinases impacts downstream signaling and drug binding. *Mol Cell* **43**, 723-737.
- Krautler, V., van Gasteren, W.F., Huneberger, P.H. (2001) A fast SHAKE algorithm to solve distant constraint equations for small molecules in molecular dynamics simulations. *J Comp Chem* **22**, 501-508.
- Kruse, A. C., Ring, A. M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hübner, H., Pardon, E., Valant, C., Sexton, P. M., Christopoulos, A., Felder, C. C., Gmeiner, P., Steyaert, J., Weis, W. I., Garcia, K. C., Wess, J., and Kobilka, B. K. (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101–106.
- Kuwabara, P. E., and Labouesse, M. (2002) The sterol-sensing domain: multiple families, a unique role? *Trends in Genetics* **18**, 193–201.
- Leahy, D.J., Dann C.E. III, Longo, P., Perman, B., Ramyar, K.X. (2000) A mammalian expression vector for purification of secreted proteins for structural studies. *Protein Expression and Purification* **20**, 500-506.
- Lee, J., Miyazaki, M., Romeo, G.R., Shoeson, S.E. (2014) Insulin receptor activation with transmembrane domain ligands. *J Biol Chem* **289**:19769-19777.
- Lemmon, M. A., and Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134.
- Lemmon, M. A., Schlessinger, J. and Ferguson, K. M. (2014) The EGFR family: Not so prototypical receptor tyrosine kinases. *Cold Spring Harb Perspect Biol* **6**, a020768.
- Lew, E.D., Furdui, C.M., Anderson, K.S., and Schlessinger, J. (2009) The precise sequence of FGF receptor autophosphorylation is kinetically driven and is disrupted by oncogenic mutations. *Sci Signal* **2**, ra6.
- Lewinson, O., Lee, A. T., and Rees, D. C. (2008) The funnel approach to the precrystallization production of membrane proteins. *J Mol Biol* **377**, 62-73.

- Li, E., Placone, J., Merzlyakov, M., and Hristova, K. (2008) Quantitative measurements of protein interactions in a crowded cellular environment. *Anal Chem* **80**, 5976-5985.
- Lippert R.A., Bowers, K.J., Gregersen, B.A., Dror, R.O, Eastwood, M.P., Klepeis, J.L, Klossvary, I., and Shaw, D.E. (2007) A common, avoidable source of error in molecular dynamics integrators. *J Chem Phys* **126**, 046101.
- Liu, J.P., Baker, J., Perkins, A.S., Robertson, E.J., and Efstratiadis, A. (1993) Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). *Cell* **75**, 59-72.
- Liu, P., Cleveland, T.E. IV, Bouyain S., Byrne, P.O., Longo, P.A., and Leahy D.J. (2012) A single ligand is sufficient to activate EGFR dimers. *Proc Natl Acad Sci USA* **109**, 10861-10866.
- Lu, X., Liu, S., & Kornberg, T. B. (2006) The C-terminal tail of the Hedgehog receptor Patched regulates both localization and turnover. *Genes Dev* **20**, 2539-51.
- Mackerell, A.D., Jr., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586-3616.
- Marigo, V., Davey, R. A., Zuo, Y., Cunningham, J. M., and Tabin, C. J. (1996) Biochemical evidence that patched is the Hedgehog receptor. *Nature* **384**, 176-9.
- Marigo, V., Scott, M. P., Johnson, R. L., Goodrich, L. V., & Tabin, C. J. (1996). Conservation in hedgehog signaling: induction of a chicken patched homolog by Sonic hedgehog in the developing limb. *Development* **122**, 1225-33.
- Mayer, S.C., Banker, A.L, Boschelli, F., Di, L., Johnson, M., Kenny, C.H., Krishnamurthy, G., Kutterer, K. Moy, F., Ptusky, S., Ravi, M., Tkach, D., Tsour, H.R., Xu, W. (2008) Lead identification to generate isoquinolinedione inhibitors of insulin-like growth factor receptor (IGF-1R) for potential use in cancer treatment. *Bio Med Chem Letters* **18**, 3641-3645.
- McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-5.
- McKern, N.M., Lawrence, M.C., Streltsov, V.A., Lou, M.Z., Adams, T.E., Lovrecz, G.O., Elleman, T.C., Richards, K.M., Bentley, J.D., Pilling, P.A., et al. (2006) Structure of the insulin receptor ectodomain reveals a folded-over conformation. *Nature* **443**, 218-221.



- McLellan, J. S., Yao, S., Zheng, X., Geisbrecht, B. V., Ghirlando, R., Beachy, P. A., and Leahy, D. J. (2006) Structure of a heparin-dependent complex of Hedgehog and Ihog. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17208–17213.
- McLellan, J. S., Zheng, X., Hauk, G., Ghirlando, R., Beachy, P. A., and Leahy, D. J. (2008) The mode of Hedgehog binding to Ihog homologues is not conserved across different phyla. *Nature* **455**, 979–983.
- Menting, J.G., Whittaker, J., Margetts, M.B., Whittaker, L.J., Kong, G.K., Smith, B.J., Watson, C.J., Zakova, L., Kletvikova, E., Jiracek, J., et al. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241–245.
- Munshi, S., Kornienko, M., Hall, D.L., Reid, J.C., Waxman, L., Stirdivant, S.M., Darke, P.L., and Kuo, L.C. (2002). Crystal structure of the Apo, unactivated insulin-like growth factor-1 receptor kinase. Implication for inhibitor specificity. *J Biol Chem* **277**, 38797–38802.
- Myers, B. R., Sever, N., Chong, Y. C., Kim, J., Belani, J. D., Rychnovsky, S., Bazan, J. F., and Beachy, P. A. (2013) Hedgehog Pathway Modulation by Multiple Lipid Binding Sites on the Smoothened Effector of Signal Response. *Dev. Cell* **26**, 346–357.
- Mynarcik, D.C., Yu, G.Q., and Whittaker, J. (1996) Alanine-scanning mutagenesis of a C-terminal ligand binding domain of the insulin receptor alpha subunit. *J Biol Chem* **271**, 2439–2442.
- Nachtergaele, S., Mydock, L. K., Krishnan, K., Rammohan, J., Schlesinger, P. H., Covey, D. F., and Rohatgi, R. (2012) Oxysterols are allosteric activators of the oncoprotein Smoothened. *Nature Chem. Biol.* **8**, 211–220.
- Nachtergaele, S., Whalen, D. M., Mydock, L. K., Zhao, Z., Malinauskas, T., Krishnan, K., Ingham, P. W., Covey, D. F., Siebold, C., and Rohatgi, R. (2013) Structure and function of the Smoothened extracellular domain in vertebrate Hedgehog signaling. *Elife* **2**, e.01340.
- Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J. R., and Ingham, P. W. (1989) A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene patched. *Nature* **341**, 508–513.
- Nedelcu, D., Liu, J., Xu, Y., Jao, C., and Salic, A. (2013) Oxysterol binding to the extracellular domain of Smoothened in Hedgehog signaling. *Nature Chem. Biol.* **9**, 557–564.
- Neria, E., Rischer, S., Karplus, M. (1996) Simulation of activation free energies in molecular systems. *J Chem Phys* **105**, 1902–1921.

- Nguyen, T. A., Lieu, S. S., and Chang, G. (2010) An Escherichia coli-based cell-free system for large-scale production of functional mammalian membrane proteins suitable for X-ray crystallography. *J Mol Micr Biot* **18**, 85-91.
- Nüsslein-Volhard, C., and Wieschaus, E. (1980) Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795–801.
- Nygaard, R., Zou, Y., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. J., Bokoch, M. P., Thian, F. S., Kobilka, T. S., Shaw, D. E., Mueller, L., Prosser, R. S., and Kobilka, B. K. (2013) The Dynamic Process of  $\beta$ 2-Adrenergic Receptor Activation. *Cell* **152**, 532–542.
- Parang, K., Till, J.H., Ablooglu, A.J, Kohanski, R.A., Hubbard, S.R., and Cole, P.A. (2001) Mechanism-based design of a protein kinase inhibitor. *Nat Struct Biol* **8**, 37-41.
- Pautsch, A., Zoephel, A., Ahorn, H., Spevak, W., Hauptmann, R., and Nar, H. (2001) Crystal structure of bisphosphorylated IGF-1 receptor kinase: insight into domain movements upon kinase activation. *Structure* **9**, 955-965.
- Philipp, M., and Caron, M. G. (2009) Hedgehog signaling: is Smo a G protein-coupled receptor? *Curr. Biol.* **19**, R125–7.
- Piana, S., Lindorff-Larsen, K., and Shaw, D.E. (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophys J* **100**, L47-49.
- Pike, L.J., Eakes, A.T., and Krebs, E.G. (1986) Characterization of affinity-purified insulin receptor/kinase. Effects of dithiothreitol on receptor/kinase function. *J Biol Chem* **261**, 3782-3789.
- Polizio, A. H., Chinchilla, P., Chen, X., Kim, S., Manning, D. R., and Riobo, N. A. (2011) Heterotrimeric Gi proteins link Hedgehog signaling to activation of Rho small GTPases to promote fibroblast migration. *Biol Chem* **286**, 19589–19596.
- Polizio, A. H., Chinchilla, P., Chen, X., Manning, D. R., and Riobo, N. A. (2011) Sonic Hedgehog activates the GTPases Rac1 and RhoA in a Gli-independent manner through coupling of smoothened to Gi proteins. *Sci Signal* **4**, pt7.
- Pollak, M. (2012) The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat Rev Cancer* **12**, 159-169.
- Qiu, C., Tarrant, M.K., Boronina, T., Longo, P.A., Kavran, J.M., Cole, R.N., Cole, P.A., and Leahy, D.J. (2009) In vitro enzymatic characterization of near full length EGFR in activated and inhibited states. *Biochemistry* **48**, 6624-6632.

- Radhakrishnan, A., Sun, L.-P., Kwon, H. J., Brown, M. S., and Goldstein, J. L. (2004) Direct binding of cholesterol to the purified membrane region of SCAP: mechanism for a sterol-sensing domain. *Mol Cell* **15**, 259-68.
- Raicu, V., Stoneman, M.R., Fung, R., Melnichuk, M., Jansma, D.B., Pisterzi, L.F., Rath, S., Fox, M., Wells, J.W., and Saldin, D.K. (2008) Determination of supramolecular structure and spatial distribution of protein complexes in living cells. *Nat Photonics* **3**, 107-113.
- Rana, R., Carroll, C. E., Lee, H.-J., Bao, J., Marada, S., Grace, C. R. R., Guibao, C. D., Ogden, S. K., and Zheng, J. J. (2013) Structural insights into the role of Smoothened cystein-rich domain in Hedgehog signaling. *Nat Commun* **4**, 2965.
- Ribes, V., and Briscoe, J. (2009) Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. *Cold Spring Harb Perspect Biol* **1**, a002014.
- Riddle, R. L., Johnson, R. L., Laufer, E., and Tabin, C. (1993) Sonic-hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401-1416.
- Riedel, H., Dull, T. J., Schlessinger, J., and Ullrich, A. (1986) A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor. *Nature* **324**, 68-70.
- Riedel, H., Dull, T. J., Honegger, A. M., Schlessinger, J. and Ullrich, A. (1989) Cytoplasmic domains determine signal specificity, cellular routing characteristics and influence ligand binding of epidermal growth factor and insulin receptors. *EMBO J* **10**, 2943-2954.
- Ring, A. M., Manglik, A., Kruse, A. C., Enos, M. D., Weis, W. I., Garcia, K. C., and Kobilka, B. K. (2013) Adrenaline-activated structure of  $\beta$ 2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575–579.
- Riobo, N. A., Saucy, B., DiLizio, C., and Manning, D. R. (2006) Activation of heterotrimeric G proteins by Smoothened. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12607–12612.
- Robarge, K. D., Brunton, S. A., Castanedo, G. M., Cui, Y., Dina, M. S., Goldsmith, R., Gould, S. E., Guichert, O., Gunzner, J. L., Halladay, J., Jia, W., Khojasteh, C., Koehler, M. F. T., Kotkow, K., La, H., Lalonde, R. L., Lau, K., Lee, L., Marshall, D., Marsters, J. C., Murray, L. J., Qian, C., Rubin, L. L., Salphati, L., Stanley, M. S., Stibbard, J. H. A., Sutherlin, D. P., Ubhayaker, S., Wang, S., Wong, S., and Xie, M. (2009) GDC-0449-a potent inhibitor of the hedgehog pathway. *Bioorg. Med. Chem. Lett.* **19**, 5576–5581
- Rohatgi, R., Milenkovic, L., and Scott, M. P. (2007) Patched1 Regulates Hedgehog Signaling at the Primary Cilium. *Science* **317**, 372–376.

Rohatgi, R., Milenkovic, L., Corcoran, R. B., and Scott, M. P. (2009) Hedgehog signal transduction by Smoothened: Pharmacologic evidence for a 2-step activation process. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3196–3201.

Rosenbaum, D. M., Rasmussen, S. G. F., and Kobilka, B. K. (2009) The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363.

Russ, W.P., and Engelman, D.M. (2000). The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol* 296, 911-919.

Ryan, K. E., and Chiang, C. (2012) Hedgehog Secretion and Signal Transduction in Vertebrates. *J. Biol. Chem.* **287**, 17905–17913

Schaffer, L. (1994) A model for insulin binding to the insulin receptor. *Eur J Biochem* **221**, 1127-1132.

Schaffer, L., and Ljungqvist, L. (1992) Identification of a disulfide bridge connecting the alpha-subunits of the extracellular domain of the insulin receptor. *Biochem Biophys Res Commun* **189**, 650-653.

Shan Y., Klepeis, J.L., Eastwood, M.P, Dror, R.O, and Shaw, D.E. (2005) Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. *J Chem Phys* **122**, 054101.

Shaw, D.E., Dror, R.O., Salmon, J.K., Grossman, J.P., Mackenzie, K.M., Bank, J.A., Young, C., Deneroff, M.M., Batson, B., Bowers, K.J., et al. (2009). Millisecond-scale molecular dynamics simulations on Anton. *Proceedings of the Conference on High Performance Computing, Networking, Storage, and Analysis* **39**, 1-11.

Shen, F., Cheng, L., Douglas, A. E., Riobo, N. A., and Manning, D. R. (2013) Smoothened is a fully competent activator of the heterotrimeric G protein G(i). *Mol. Pharmacol.* **83**, 691–697.

Shoelson, S.E., White, M.F., and Kahn, C.R. (1988) Tryptic activation of the insulin receptor. Proteolytic truncation of the alpha-subunit releases the beta-subunit from inhibitory control. *J Biol Chem* **263**, 4852-4860.

Siddle, K. (2011) Signalling by insulin and IGF receptors: supporting acts and new players. *J Mol Endocrinol* **47**, R1-10.

Siddle, K. (2012) Molecular basis of signaling specificity of insulin and IGF receptors: neglected corners and recent advances. *Front Endocrinol (Lausanne)* **3**, 34.

Smith, B.J., Huang, K., Kong, G., Chan, S.J., Nakagawa, S., Menting, J.G., Hu, S.Q., Whittaker, J., Steiner, D.F., Katsoyannis, P.G., et al. (2010) Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. *Proc Natl Acad Sci U S A* **107**, 6771-6776.

Songyang, Z., Carraway, K.L., 3rd, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., et al. (1995) Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536-539.

Soos, M.A., Field, C.E., Lammers, R., Ullrich, A., Zhang, B., Roth, R.A., Andersen, A.S., Kjeldsen, T., and Siddle, K. (1992) A panel of monoclonal antibodies for the type 1 insulin-like growth factor receptor. Epitope mapping, effects on ligand binding, and biological activity. *J Biol Chem* **267**, 12955-12963.

Sparrow, L. G., McKern, N. M., Gorman, J. J., Strike, P. M., Robinson, C. P., Bentley, J. D., and Ward, C. W. (1997) The disulfide bonds in the C-terminal domains of the human insulin receptor ectodomain. *J Biol Chem* **272**, 29460-29467.

Stone, D. M., Hynes, M., Armanini, M., Swanson, T. A., Gu, Q., Johnson, R. L., Scott, M. P., Pennica, D., Goddard, A., Phillips, H., Noll, M., Hooper, J. E., de Sauvage, F., and Rosenthal, A. (1996) The tumour-suppressor gene patched encodes a candidate receptor for Sonic hedgehog. *Nature* **384**, 129–134.

Sweeney, R. T., McClary, A. C., Myers, B. R., Biscocho, J., Neahring, L., Kwei, K. A., Qu, K., Gong, X., Ng, T., Jones, C. D., Varma, S., Odegaard, J. I., Sugiyama, T., Koyota, S., Rubin, B. P., Troxell, M. L., Pelham, R. J., Zehnder, J. L., Beachy, P. A., Pollack, J. R., and West, R. B. (2014) Identification of recurrent SMO and BRAF mutations in ameloblastomas. *Nat. Genet.* **46**, 722–725.

Taipale, J., Chen, J. K., Cooper, M. K., Wang, B., Mann, R. K., Milenkovic, L., Scott, M. P., and Beachy, P. A. (2000) Effects of oncogenic mutations in Smoothened and Patched can be reversed by cyclopamine. *Nature* **406**, 1005–1009.

Taipale, J., Cooper, M. K., Maiti, T., and Beachy, P. A. (2002) Patched acts catalytically to suppress the activity of Smoothened. *Nature* **418**, 892–897.

Tamura, S., Fujita-Yamaguchi, Y., and Larner, J. (1983) Insulin-like effect of trypsin on the phosphorylation of rat adipocyte insulin receptor. *J Biol Chem* **258**, 14749-14752.

Therond, P. P. (2012) Release and transportation of Hedgehog molecules. *Curr. Opin. Cell. Biol.* **24**, 173–180.

Till, J.H., Ablooglu, A.J, Frankel, M., Bishop, S.M., Kohanski, R.A., and Hubbard, S.R. (2001) Crystallographic and solution studies of an activation loop mutant of the insulin receptor tyrosine kinase: insights into kinase mechanism. *J Biol Chem* **276**:10049-10055.

Tseng, T. T., Gratwick, K. S., Kollman, J., Park, D., Nies, D. H., Goffeau, A., and Saier, M. H. (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family

that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* **1**, 107–125.

Tuckerman, M., Berne, B.J. (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* **97**, 1990-2001.

Ullrich, A., Bell, J.R., Chen, E.Y., Herrera, R., Petruzzelli, L.M., Dull, T.J., Gray, A., Coussens, L., Liao, Y.C., Tsubokawa, M., et al. (1985). Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* **313**, 756-761.

van den Heuvel, M., and Ingham, P. W. (1996) smoothened encodes a receptor-like serpentine protein required for hedgehog signalling. *Nature* **382**, 547–551.

van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008) Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **9**, 112-124.

Veleparthi, U., Wittman, M., Liu, P., Stoffan, K. Zimmermann, K., Sang, X., Carboni, J., Li, A., Attar, R., Gottardis, M., Greer, A., Chang, C.Y., Jacobsen, B.L, Sack, J.S., Sun, Y., Langley, D.R., Balasubramanian, B., and Vyas, D. (2007) Discovery and initial SAR of 3-(1H,benzol[f]imidazol-2-yl)pyridin-2(1H)-ones as inhibitors of insulin-like growth factor 1-receptor (IGF-1R). *Bio Med Chem Letters* **17**, 2317-2321.

Venkatakrishnan, A. J., Deupi, X., Lebon, G., Tate, C. G., Schertler, G. F., and Babu, M. M. (2013) Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194.

Wang, C., Wu, H., Evron, T., Vardy, E., Han, G. W., Huang, X.-P., Hufeisen, S. J., Mangano, T. J., Urban, D. J., Katritch, V., Cherezov, V., Caron, M. G., Roth, B. L., and Stevens, R. C. (2014) Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. *Nat Commun* **5**, 4355.

Wang, C., Wu, H., Katritch, V., Han, G. W., Huang, X.-P., Liu, W., Siu, F. Y., Roth, B. L., Cherezov, V., and Stevens, R. C. (2013) Structure of the human smoothened receptor bound to an antitumour agent. *Nature* **497**, 338–343.

Wang, W., Black, S. S., Edwards, M. D., Miller, S., Morrison, E. L., Bartlett, W., et al. (2008) The structure of an open form of an E. coli mechanosensitive channel at 3.45 Å resolution. *Science* **321**, 1179-83.

Wang, Y., Davidow, L., Arvanites, A. C., Blanchard, J., Lam, K., Xu, K., Oza, V., Yoo, J. W., Ng, J. M. Y., Curran, T., Rubin, L. L., and McMahon, A. P. (2012) Glucocorticoid compounds modify smoothened localization and hedgehog pathway activity. *Chem. Biol.* **19**, 972–982.

Ward, C.W., and Lawrence, M.C. (2012). Similar but different: ligand-induced activation of the insulin and epidermal growth factor receptor families. *Curr Opin Struct Biol* **22**, 360-366.

Ward, C.W., Menting, J.G., and Lawrence, M.C. (2013). The insulin receptor changes conformation in unforeseen ways on ligand binding: sharpening the picture of insulin receptor activation. *Bioessays* **35**, 945-954.

Warne, T., Moukhametzianov, R., Baker, J. G., Nehme, R., Edwards, P. C., Leslie, A. G. W., Schertler, G. F. X., and Tate, C. G. (2011) The structural basis for agonist and partial agonist action on a B1-adrenergic receptor. *Nature* **469**, 241–245.

Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Doak, R. B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Han, G. W., Katritch, V., Boutet, S. E. B., Messerschmidt, M., Williams, G. J., Koglin, J. E., Seibert, M. M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M., and Cherezov, V. (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat Commun* **5**, 3309.

Whittaker, J., Groth, A. V., Mynarcik, D. C., Pluzek, L., Gadsboll, V. L., and Whittaker, L. J. (2001) Alanine scanning mutagenesis of a type 1 insulin-like growth factor receptor ligand binding site. *J Biol Chem* **276**, 43980-43986.

Whittaker, L., Hao, C., Fu, W., and Whittaker, J. (2008) High-affinity insulin binding: insulin interacts with two receptor ligand binding sites. *Biochemistry* **47**, 12900-12909.

Williams, E. H., Pappano, W. N., Saunders, A. M., Kim, M.-S., Leahy, D. J., and Beachy, P. A. (2010) Dally-like core protein and its mammalian homologues mediate stimulatory and inhibitory effects on Hedgehog signal response. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5869–5874

Williams, P.F., Mynarcik, D.C., Yu, G.Q., and Whittaker, J. (1995) Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. *J Biol Chem* **270**, 3012-3016.

Wu, J., Li, W., Craddock, B. P., Foreman, K.W., Mulvihill, M. J., Ji, Q. S., Miller, W. T., and Hubbard, S. R. (2008) Small-molecule inhibition and activation-loop trans-phosphorylation of the IGF1 receptor. *EMBO J* **27**, 1985-1994.

Xie, J., Murone, M., Luoh, S.-M., Ryan, A., Gu, Q., Zhang, C., Bonifas, J. M., Lam, C.-W., Hynes, M., Goddard, A., Rosenthal, A., Jr, E. H. E., and de Sauvage, F. J. (1998) Activating Smoothed mutations in sporadic basal-cell carcinoma. *Nature* **391**, 90–92.

Yamada, K., Goncalves, E., Kahn, C.R., and Shoelson, S.E. (1992) Substitution of the insulin receptor transmembrane domain with the c-neu/erbB2 transmembrane domain constitutively activates the insulin receptor kinase in vitro. *J Biol Chem* **294**, 1-14.

Yao, S., Lum, L., and Beachy, P. (2006) The ihog cell-surface proteins bind Hedgehog and mediate pathway activation. *Cell* **125**, 343–357.

Yauch, R. L., Dijkgraaf, G. J. P., Alicke, B., Januario, T., Ahn, C. P., Holcomb, T., Pujara, K., Stinson, J., Callahan, C. A., Tang, T., Bazan, J. F., Kan, Z., Seshagiri, S., Hann, C. L., Gould, S. E., Low, J. A., Rudin, C. M., and de Sauvage, F. J. (2009) Smoothed Mutation Confers Resistance to a Hedgehog Pathway Inhibitor in Medulloblastoma. *Science* **326**, 572–574.

Zachowski, A. (1993) Phospholipids in animal eukaryotic membranes: transverse asymmetry and movement. *Biochem J* **294**, 1-14.

Zhang, C., Williams, E. H., Guo, Y., Lum, L., and Beachy, P. A. (2004) Extensive phosphorylation of Smoothed in Hedgehog pathway activation. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17900–17907.

Zhang, X., Gureasko, J., Shen, K., Cole, P.A., and Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* **125**, 1137-1149.

Zhao, Y., Tong, C., and Jiang, J. (2007) Hedgehog regulates smoothed activity by inducing a conformational switch. *Nature* **450**, 252–258.

Zheng, X., Mann, R. K., Sever, N., and Beachy, P. A. (2010) Genetic and biochemical definition of the Hedgehog receptor. *Genes Dev.* **24**, 57–71



## *Curriculum Vitae*

### **JACQUELINE MELLE MCCABE**

#### PLACE OF BIRTH

Wynnewood, Pennsylvania

#### EDUCATION

##### **Johns Hopkins University School of Medicine**

Baltimore, MD

PhD (candidate): Biophysics and Biophysical Chemistry

August 2010 - July 2015

Advisor: Daniel J. Leahy, PhD

Dissertation: *How IGF1 activates its receptor fails to activate EGFR*

##### **Georgetown University**

Washington, D.C.

Bachelor of Science with Honors

August 2006 - May 2010

Major: Biochemistry

*Magna Cum Laude*

#### RESEARCH EXPERIENCE

##### **Johns Hopkins University School of Medicine**

Baltimore, MD

Graduate Research

August 2010 - July 2015

##### **Georgetown University**

Washington, D.C.

Undergraduate Research

August 2007 - May 2010

Advisor: Paul D. Roepe, PhD

Honors Thesis: *Kinetic analysis of chloroquine transport in malaria parasites*

#### TEACHING EXPERIENCE

##### **Teaching Assistant and Group Discussion Leader**

October 2012

Macromolecules Section, Scientific Foundations of Medicine Course

##### **Biophysics Tutorial Leader**

2011 - 2014

Biochemical and Biophysical Principles, Graduate Student Core Course

#### FELLOWSHIPS

##### **National Science Foundation Graduate Research Fellow**

2012 - 2015

##### **GUROP Summer Research Fellowship**

2008, 2009

##### **Chemistry Department Adams Research Fellowship**

2009

#### AWARDS

##### **Second Place in Poster Presentation, BCMB Retreat**

2012

##### **Award for Outstanding Undergraduate Research in Biochemistry**

2010

##### **American Institute of Chemists Student Award**

2010

##### **Chemistry Department Award for Distinction in Chemistry**

2010

##### **Dean's List, First Honors and Second Honors**

2006 - 2010

### EDUCATIONAL OUTREACH

**Mentor and College Prep Committee Member** 2010 - 2013  
Incentive Mentoring Program (THREAD)

**Alumni Interviewer** 2010 - 2014  
Georgetown University Alumni Admissions Committee

**Math and Science Tutor** 2007- 2009  
GU Math and Science Hands-On Enrichment Program

### POSTER PRESENTATION

Cleveland T. E. IV, **McCabe J. M.**, and Leahy, D. J. (2014) Detergent-solubilized and Purified Patched Fails to Interact Strongly with Cognate Hedgehog or Ihog Homologs. Hedgehog 2014 Meeting. Ann Arbor, Michigan.

### PUBLICATIONS

Cleveland, T. E. IV, McCabe J. M., and Leahy, D. J. (2014) Detergent-solubilized Patched purified from Sf9 cells fails to interact strongly with cognate Hedgehog or Ihog homologs. *Protein Expr Purif* **104C**, 92-102.

Kavran, J. M., McCabe, J. M., Byrne, P. O., Connacher, M. K., Wang, Z., Ramek, A., Sarabiopour, S., Shan, Y., Shaw, D. E., Hristova, K., Cole P. A., and Leahy, D. J. (2014) How IGF-1 activates its receptor. *Elife*, doi: 10.7554/eLife.03772.

McCabe, J. M. and Leahy, D. J. (2015) Smoothened goes molecular: new pieces in the hedgehog signaling puzzle. *J Biol Chem* **6**, 3500-7.