ABSTRACT

BARUA, DIPAK. Rule-Based Computational Modeling of Modular Signaling Protein Interactions. (Under the direction of Jason M. Haugh.)

Intracellular signal transduction pathways are comprised of complex interactions among cellular proteins and other biomolecules. The structures of signaling proteins/enzymes are often modular, with conserved domains that carry out specific interactions or catalytic functions, and their core activities are dictated through coordinated intra- and inter-molecular interactions. In collaboration with Prof. James Faeder (Computational Biology, University of Pittsburgh), we have applied a computational algorithm for generating large networks of kinetic equations based on a much smaller set of mechanistic rules. Using this rule-based approach, we have formulated kinetic models that account for the modular domain structure of specific signaling proteins, including Shp2 (Src homology-2 domain containing protein tyrosine phosphatase 2), PI3K (phosphoinositide 3-kinase) regulatory subunit, and SH2-B (a Jak2 kinase activating adaptor protein). Analysis of these models reveals the combinatorial possibilities of reactions and interactions that might occur in living cells. We propose here to extend this rule-based approach for larger pathway models through systematic reduction and integration of small subsystem models.

Rule-based Computational Modeling of Modular Signaling Protein Interactions

by
Dipak Barua

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Chemical Engineering

Raleigh, North Carolina

August 2008

APPROVED BY:

(Chair of Advisory Committee) (Jason M. Haugh)	(Orlin D. Velev)
(Alun L. Lloyd)	(Balaji M. Rao)

BIOGRAPHY

Dipak Barua was born in Bangladesh, a country in South Asia, bordered by India and Myanmar, and coasted by the Bay of Bengal. He grew up with his parents in a small port city of Bangladesh, where he completed his school education, and then moved to Dhaka (Capital of Bangladesh) to begin his college education. Dipak's long cherished goal of becoming an engineer (although in an undetermined field) came true when he was admitted in Bangladesh University of Engineering and Technology (BUET), the premier engineering institution of the country. He chose to be a chemical engineer eventually, and received his BS diploma in April 2002. In August 2003, Dipak left for US to attend graduate school in Chemical Engineering at North Carolina State University. He completed MS degree in chemical engineering in 2005, and motivated to continue in PhD thereafter.

ACKNOWLEDGEMENTS

Throughout my stay in Raleigh and my PhD education at NC State, I have been developing a memoir that is enriched of helpful events by many people in my surrounding. Firstly, I am immensely grateful to my advisor, Professor Jason M. Haugh for allowing me to join his lab, and for all the supports and guidance I have been receiving from him in every steps of my research. He had been a great teacher and mentor to me, and I have learnt from him every aspect of doing research that I consider to be valuable assets and will help me ever to progress in my future career. I am also greatly indebted to Professor James R. Faeder, who had allowed me to use BioNetGen; my research entirely relies on the implementation of this modeling software. In addition, I have gained many good suggestions and advice from Professor Faeder, which helped me identifying and solving many crucial problems in my work. I also acknowledge the support and encouragement from my committee. My thanks to all my friends, my current and previous group mates in the Haugh laboratory, who had helped me many times on many occasions, related or unrelated to research. Finally, I thank my parents and rest of my family, whose invisible support always inspires me moving forward in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Modularity of cell signaling protein structures and complexity of modeling signal transduction pathways

1.1 MOTIVATION

Traditional mathematical models of signal transduction pathways are mostly simplistic in the sense that these models usually incorporate minimal (lumped) molecules and reactions in a pathway in order to capture the essential system behaviors. These simple models, although useful in qualitative understanding of many signaling pathways, are too conservative in terms of providing quantitative and mechanistic information. In recent years, the role-played by signaling molecules and their connectivities in many signal transduction networks are well established. Yet little knowledge has been gained regarding how these signaling molecules communicate and function in concert to dictate dynamical properties of pathways and networks in cells. This type of mechanistic interpretation would necessitate prior unraveling of the molecular functional mechanisms of individual signaling molecules in a pathway. The emerging concept of detailed pathway modeling[1-7], which accounts for protein-protein interactions at the level of modular protein domains, seems be a promising avenue for delving into the submolecular events that might prevail during information propagation in cell signaling networks.

In this study, we have formulated domain-based protein interaction models that, we hope, can capture the level of detail comparable to what is accessible through experimental measurement. We have devised a computational approach to enumerate a larger set of species and reactions that may evolve in a system through the domain level transformations of the reacting protein molecules. To implement this approach, we have adopted a rule-based algorithm, which allowed us to model the reactant molecules with structural attributes of certain proteins or enzymes[8]. Our models have been aimed at elucidating the structure-function relationships of modular, multi-domain signaling proteins. We have critically characterized the input-output functions and dynamic regulation of three major signaling

enzymes: a protein tyrosine phosphatase, Shp2, and a lipid kinase, PI3K, which are important for activation of the MAP kinase and Akt signal transduction pathways, respectively, and Jak2, a tyrosine kinase required for signaling through receptors for growth hormone and other cytokines. We anticipate that such comprehensive and elemental dissection of small subsystem models, as we have presented in this study, will effectively contribute to explorations of larger network function quantitatively and mechanistically, thereby broadening the span of understanding of complex response behaviors associated with different signal transduction systems.

1.2 INTRODUCTION

1.2.1 Cell signal transduction: general overview

Cell signal transduction refers to the complex intracellular processes by which living cells respond to the extracellular stimuli/ligands (e.g., growth factors, cytokines, hormones, etc) via regulation of specific cell functions such as growth, survival, differentiation, and migration[9]. Intracellular signaling processes are customarily complex and involve numerous interactions among large numbers of molecular species (e.g., proteins, lipids, and ions). In a ligand-stimulated cell, these molecules become activated in an ordered fashion, beginning at the plasma membrane and resulting in the actuation of processes involving the cell cytoskeleton or gene transcription machineries in the nucleus. In such interaction networks, signaling molecules receive the input information from their upstream regulators and deliver them to their downstream effectors. The arrow diagram in Fig. 1.1 represents the simplified form of signaling networks that become activated upon platelet derived growth factor (PDGF) mediated cell stimulation.

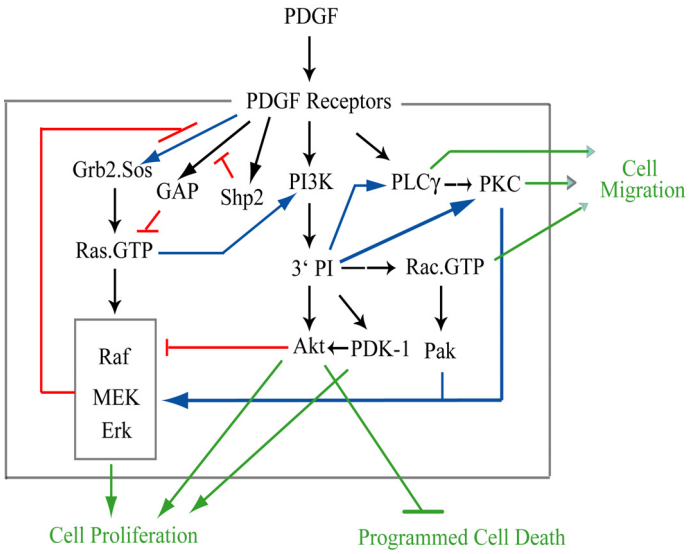


FIGURE 1.1 PDGF activated interaction network. Key signaling molecules and their connectivities in PDGF signalng pathways are shown in an arrow diagram. Figure adapted from Haugh Research Group web page (http://www.che.ncsu.edu/haughlab/)

Cell signaling pathways incorporate multiple layers of complexity. The first originates from the sheer size of a network involving a large number of signaling molecules and reactions. In addition, complexity can also arise because signaling pathways seldom

operate independently. Distinct pathways often share common species or interaction modules, such that pleiotropic effects and cross-talk interactions are typical among these pathways (Fig. 1.2). The third layer of complexity arises from the modular structure of signaling proteins/enzymes. Modular, multi-domain proteins are stringently regulated through intra-and intermolecular domain interactions. This level of complexity has thus far been discounted in most theoretical studies of signal transduction.

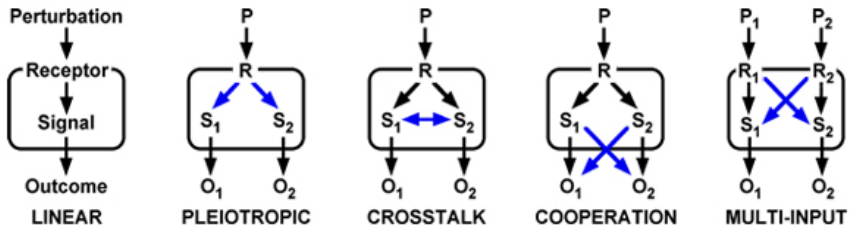


FIGURE 1.2 Interdependency of signaling pathways. The linear signal propagation is uncommon in receptor mediated signaling. Pleiotropy, crosstalk, cooperation, multi-input, or combinations of such interdependencies are typical among signaling pathways. Figure adapted from Haugh Research Group web page (http://www.che.ncsu.edu/haughlab/)

1.2.2 Structural modularity of signaling proteins and implications in signal transduction

Eukaryotic signal transduction proteins contain complex architectures built from combinations of discrete, independently folding structural units (50 to 200 amino acids) called domains that dictate their constitutive and/or signal regulated interspecies communication in an interaction network[10-12]. These modular protein structures are commonly designed to contain one core functional/catalytic domain, few regulatory/binding domains, and multiple peptides or phosphorylated peptide sequences (domain recognition motifs). Eukaryotic genome contains a limited number of different classes of domains, each of which display preferential binding/reaction propensities towards specific set of partner domains or motifs[13].

Protein domains act as the fundamental reaction elements in a signal transduction network. Modular domains and peptide motifs establish the complex information flow

circuitry by linking a signal transducer in a network with upstream effector and downstream target proteins. Regulatory domains provide complementary binding surfaces for reversible association with their recognition motifs/domains, whereas catalytic domains induce covalent modifications (e.g. phosphorylation, dephosphorylation, nucleotide exchange) of their substrate sites. The regulatory domains of an enzyme, through selectively interacting with other species, propel the core functional activity towards specific targets by discriminating among large number of putative substrates with identical reaction motifs. Additionally, the regulatory domains often dictate the core output function through complex biophysical mechanisms, e.g., protein scaffolding/colocalization, compartmental translocations, allosteric conformational change, cooperative tandem binding interactions, etc.

1.2.2a Multi-protein colocalization

Adaptor proteins mediate direct information channeling by hardwiring successive signaling components in a cell-signal network[14, 15]. The flexible modular architectures of adaptor proteins usually accommodate a variety of binding sites that colocalize species lacking complementary domains for direct mutual binding. Scaffolding proteins function in a similar fashion as the adaptors, providing docking surfaces for large multiple protein molecules[16]. Protein scaffolding directs reaction specificities and enhances reaction rates by aggregating molecules in a single complex[16, 17]. A tyrosine phosphorylated receptor tyrosine kinase (RTK) scaffolds the Src homology 2 (SH2) / phosphotyrosine binding (PTB) domain-containing molecules from cytosol, and allows them access their plasma membrane and receptor associated substrates / reaction partners (Fig. 1.3 *a*).

1.2.2b Allosteric regulation of protein/enzyme function

Many signaling enzymes display allosteric functional regulation upon interaction with their activators or target proteins[12, 18-23] (Fig. 1.3 *b*). Prominent examples include SH2 domain-containing protein tyrosine phosphatase 2 (Shp2), phosphoinositide 3-kinase (PI3K), the Src family protein tyrosine kinases, etc. These enzymes usually contain multiple regulatory domains in addition to their core functional (catalytic) domains. The regulatory

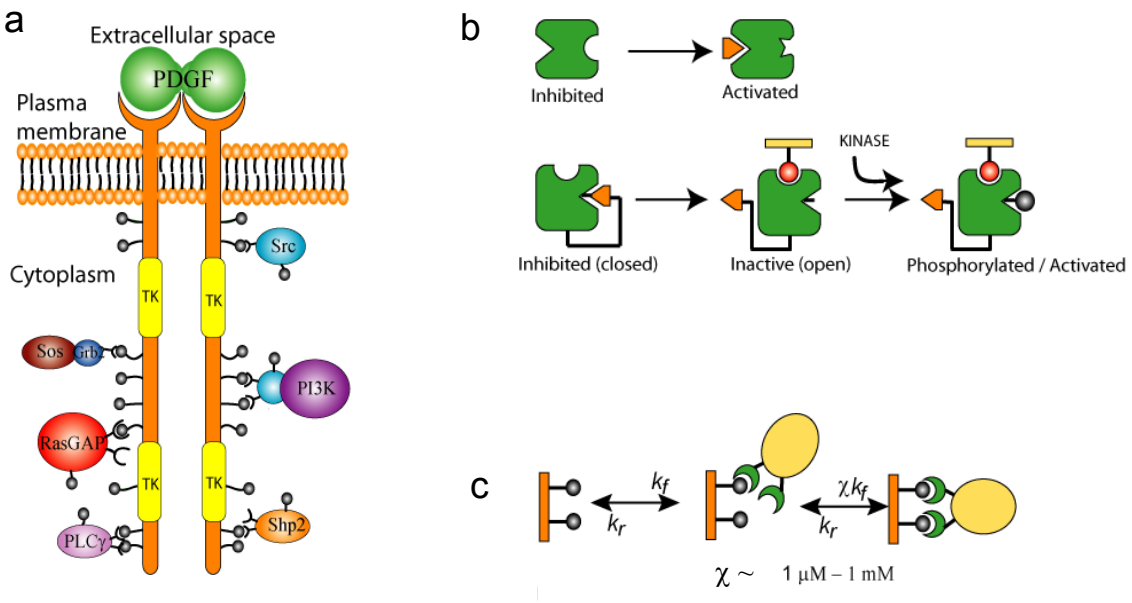


FIGURE 1.3 Examples of modular strategies of protein function regulation. (*a*) Tyrosine phosphorylated PDGF-receptors scaffold SH2 and PTB domain-containing signaling proteins. (*b*) Ligand binding induces conformational change of proteins, which activates them directly or allow them to react with an activating protein (e.g., kinase). (*c*) Cooperative interaction by multiple domains confers high affinity to protein association.

domains often incur a conformational modification in a protein's structures via associating with a binding partner, which can activate the protein by exposing its buried catalytic domain[19] and/or allowing it to interact with an activating enzyme[20, 24, 25].

1.2.2c Avidity of protein binding

Signaling proteins often utilize multiple interaction domains for associating with their binding partners[26-28]. Tandem/multidomain interactions cooperatively synergize binding affinities allowing stable protein complex formation[29, 30] (Fig. 1.4b). This type of cooperative binding directs reaction specificities by out competing low-affinity single domain protein interactions. In addition, further reaction specificity is gained due to the fact that fewer molecules present multiple complementary sites for mutual interaction[29]. For example, PI3K, ZAP-70, and Syk proteins present complementary tandem SH2 domains for specific bisphosphorylated sequences in PDGF, TCR, and FcεRI receptors, respectively.

1.2.3 Practical significance of understanding modular protein function

Structural modularity is thought to be associated with protein evolution, since addition or deletion of a single domain can generate a unique functional protein from an existing one [12]. Understanding the structure of a protein and how its modularity correlates with its function is crucially important to diagnose pathological aberrations that might arise from mutations/modifications of associated proteins domains. Signaling proteins are vitally implicated in pathways regulating cell cycle, growth, and proliferative responses. Loss of domain functionality in these proteins, either partially or fully, can disrupt a healthy regulatory circuit within the pathway, thus incurring uncontrollable cell growth/proliferation or other physiological malfunctions as seen in cancer and many lethal diseases. Understanding protein modularity could also be important for engineering chimeric protein molecules with desired functional properties[23, 31]. Finally, recombinant proteins expressing selected domains from different origins could be synthesized for targeted drug delivery purposes[32].

1.2.4 Combinatorial complexity of modular protein interactions: a major challenge of signal transduction pathway modeling

Domain interactions between modular signaling proteins result in combinatorial possibilities of distinct chemical species and states[2, 3, 5-7]. For example, a scaffold protein containing 10 tyrosine sites may assume 2^{10} distinct phosphorylation states, because each site can remain either phosphorylated or unphosphorylated (through kinase and phosphatase activities) in any combination. In addition, the scaffold can recruit a variety of signaling proteins in various combinations, leading to a proliferation of the potential chemical species and states. A practical example would be the T cell receptor (TCR) complex, which provides sixteen potential phosphotyrosine sites for binding of ZAP-70 (a Src family tyrosine kinase) [28, 33-36]. The ZAP-70 protein uses two distinct SH2 domains for binding phosphorylated TCR. In addition, ZAP-70 itself can remain phosphorylated or unphosphorylated on multiple tyrosine sites. The worst scenario of combinatorial complexity results from ‘runaway polymerization’ interactions, where modular species can hypothetically assemble to generate

indefinitely long chain or ring structures[37]. A typical example would be the binding of a biphosphorylated peptide with a protein with tandem SH2 domains in solution[38]. Practically, such combinatorial explosion of large number of species is not feasible, because aggregation would be restricted automatically based on possible random encounters of reacting species and their interaction affinities. Also, the theoretically possible number of species may often surpass the total amount of existing proteins in a cell. However, in reality, the numbers of molecular species and states can still be enormously large, and each of them might contribute differentially to the ultimate response of a cell. Without quantitative analysis, it is not possible to predict the possible concentrations and contributions of individual chemical species.

Combinatorial complexity imposes a major challenge for signal transduction pathway modeling due to the enormously large number of kinetic equations that might need to be solved for balancing all possible species in a network. Traditional modeling grossly avoids this complexity, whereas detailed modeling confronts the computational challenge. Recently, efforts have been focused to handle combinatorial complexities by enumerating only macro/meso-states of components when permissible [5, 39, 40]. Efforts have also been made to employ stochastic approaches for systems with small numbers of molecules [6, 37].

1.2.5 Detailed mechanistic modeling and combinatorial complexity

A detailed mathematical model is expected to incorporate the submolecular/domain level transformations and activities of signaling proteins, and hence their combinatorially possible microstates/species in an interaction network. In principle, the molecular domains and motifs, not the molecule as a whole, serve as the fundamental reaction elements in signaling pathways[2, 41], thus making these networks far more complex than what usually is represented in traditional reaction models. Most pathway models thus far remain simplistic in the sense that these models incorporate the lumped or reduced structures of pathways, and restrict model resolutions up to the molecule level[42-48]. Even the most rigorous traditional models that might have included most molecules in a pathway[49] do not describe the detailed features of a network, since these models regard signaling components as ideal

reaction elements, discounting their potentially complex structural regulation. Modular signaling proteins and enzymes often function as sophisticated information-gating elements, whose functionalities are dynamically regulated through complex intra-and/or intermolecular domain interactions[31, 50]. Due to complex domain-level regulations, the function of signaling enzymes often deviate from the classical Michaelis-Menten kinetics, exhibiting non-linear behaviors[31, 51]. In addition, domain- mediated binding affinities, specificities, and colocalizations contribute significantly to determine the net performance of a modular enzyme.

1.2.6 Rule-based modeling: an approach to circumvent combinatorial complexity

Rule-based mathematical modeling is a relatively new method for detailed biochemical pathway modeling that has grown in popularity because of its applicability to problems with combinatorial complexity[52]. Rule-based modeling allows one to define theoretical protein molecules with structural and functional attributes, such as sub-molecular protein components (interaction domains, motifs and subunits), and their activity states. Traditional ways of devising mathematical models require manual specification of every pertinent chemical species and reaction in a system. In contrast, rule-based construction of mathematical models involves automatic generation of the species and reactions in a network through the implementation of user-defined generic rules[37]. Each specified rule corresponds to a class of biophysical transformations associated with a particular protein-component rather than the entire protein. A modeler defines only the rudimentary or 'seed' species, which are initially present in a system prior to the network activation, and the reaction rules are applied iteratively over these seed species and any species generated in past iterations, until the process generates the entire set of combinatorial species and reactions. The exhaustive iterative calculation is performed through the graph theoretic method described elsewhere by Blinov et al (Blinov et al, Book Chapter 2006).

1.2.7 Modular proteins of interest

1.2.7a Src homology 2 domain-containing protein tyrosine phosphatase 2 (Shp2)

Shp2 is a non-transmembrane protein tyrosine phosphatase ubiquitously expressed in all mammalian cell types. Shp2 contributes positively to the activation of Ras/Erk signaling pathway mediated by platelet-derived growth factor (PDGF) receptors, epidermal growth factor (EGF) receptors, and other receptor tyrosine kinases[53-56]. The enzyme contains a complex modular structure comprised of two SH2 domains (N-SH2 and C-SH2), a centrally located catalytic (PTP) domain, and an extended C-terminal tail[57]. The C-terminal tail contains two tyrosine sites that are subject to phosphorylation by receptor tyrosine

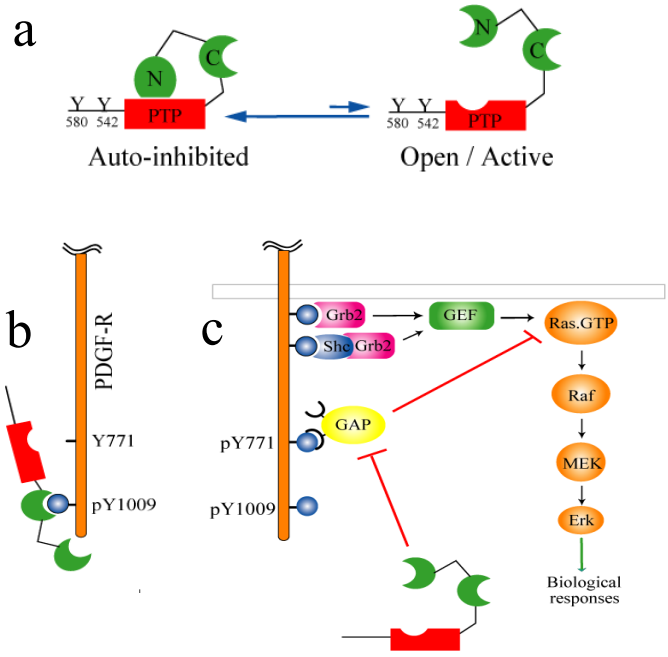


FIGURE 1.4 Shp2 regulation of PDGF signaling. (a) The N-SH2 domain of Shp2 intramolecularly binds to its PTP domain leading to an auto-inhibited conformation of the enzyme. In basal/cytoplasmic state, equilibrium is highly favored towards this closed/inactive state of Shp2. (b) Shp2 binds with Tyr^{1009} of PDGF-R by either of the SH2 domains, and dephosphorylates Tyr^{771} by the PTP domain. (c) Shp2 mediated dephosphorylation of Tyr^{771} inhibits PDGF-R recruitment/activation of the Ras GTPase activating protein, GAP, which is a negative regulator of the MAPK/Erk signaling pathway.

kinases[58-60]. In the basal condition, the N-SH2 domain of Shp2 interacts intramolecularly with the PTP domain, thereby blocking it from accessing substrates [57, 61-63] (Fig. 1.4 *a*). There is also evidence that the SH2 domains of Shp2 intramolecularly bind to the two phosphorylated tyrosines in its carboxyl terminal tail[58, 59, 64-66]; apparently, the N-SH2 domain interaction with these phosphotyrosines disrupt the auto-inhibition, favoring the open/active conformation of Shp2.

Shp2, via its two SH2 domains, binds to phosphorylated Tyr^{1009} of PDGFβ-receptors[67](Fig. 1.4 *b*). The putative substrate of Shp2 on PDGFβ-receptor is Tyr^{771}, which is also the binding/activation site for Ras GTPase- activating protein (RasGAP)[68-70]. RasGAP is a negative regulator of the Ras/Erk pathway; therefore, by dephosphorylating Tyr^{771}, Shp2 indirectly imparts a positive effect on the Ras/Erk signaling pathway (Fig. 1.4 *c*). Shp2 apparently follows complex dynamic regulation of catalytic activity by binding with RTKs. Although certain structure-function relationships for Shp2 have been elucidated to a significant extent *in vitro*, the mechanisms by which Shp2 functions in cells remains speculative.

1.2.7b Phosphoinositide 3-kinase (PI3K)

PI3K is a lipid kinase and a crucial mediator of cell survival, proliferation, and migration pathways[71-74]. The Class IA PI3Ks are heterodimeric complexes comprised of a 85 kDa regulatory subunit and a 110-kDa catalytic subunit[75, 76] (Fig. 1.5). Multi-domain structure of the p85 regulatory unit strongly activates the kinase through RTK interactions. The p85 protein contains two tandem SH2 domains (C-SH2 and N-SH2) that play the crucial role in

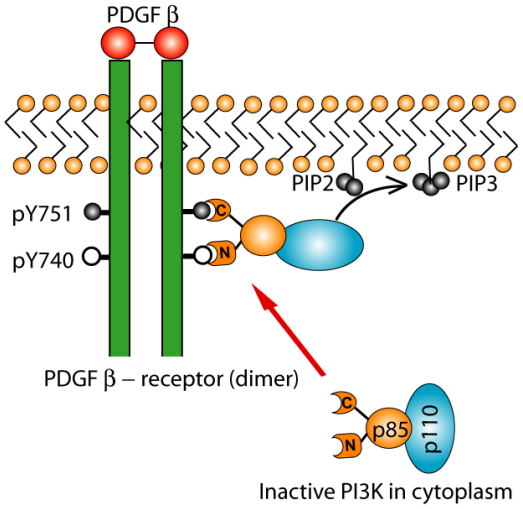


FIGURE 1.5 PI3K – PDGF-receptor interaction. Tandem binding of the two P85 SH2 domains with a bisphosphorylated PDGF receptor sequence induces highly stable receptor – PI3K complex formation.

target and allosterically activate the enzyme in cells[26, 77-79]. The N-SH2 domain mediates a conformation change in p85 upon binding with receptor phosphotyrosines, thus exposing a tyrosine site (Tyr^{688}), that is buried in the basal state[77, 79, 80]. The exposed tyrosine subsequently becomes phosphorylated by membrane associated tyrosine kinases, Lck and Abl, which impart the full catalytic potency to the enzyme[25]. Despite this important role of the N-SH2 domain in PI3K activation, its affinity for RTK phosphotyrosine binding is too low to mediate stable complex formation [81]. Apparently, both SH2 domains of PI3K contribute to the high affinity receptor association through cooperative tandem binding[29, 30]. In addition to the RTK-mediated activation, PI3K might be activated via an intramolecular N-SH2 domain – phosphotyrosine interaction[18, 25].

1.2.7c SH2 domain containing adaptor protein, SH2-B

The SH2-Bβ is a cytosolic adaptor protein, and a positive regulator of Janus kinase 2 (Jak2)-mediated growth hormone (GH) signaling[82-84]. Upon growth hormone stimulation, the adaptor directly associates with Jak2, and stimulates the kinase phosphorylation/activation through some unknown mechanism. SH2-B contains a C-terminal SH2 domain, a centrally located pleckstrin homology (PH) domain and a N-terminal dimerization domain [85, 86]. The protein forms homodimers through its dimerization domain-mediated interactions[85, 86]. SH2-B also associates with Jak2 via its SH2 domain binding with phosphorylated Tyr^{813} of Jak2 [83, 85]. In addition, SH2-Bβ associates with plasma membrane phosphoinositides (PIs) via its PH domain [85, 87].

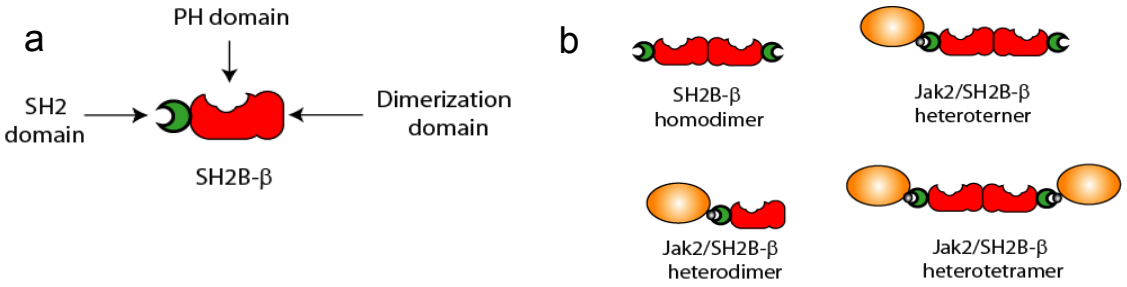


FIGURE 1.6 SH2-Bβ adaptor protein. (*a*) Three major functional domains of SH2-Bβ adaptor protein. (*b*) SH2-Bβ homodimerization and interaction with Jak2.

Although domain interactions between SH2-B and Jak2 are well known, exactly how the adaptor stimulates Jak2 phosphorylation has remained poorly understood. Based on the known interactions, a conceptual heterotetrameric complex, Jak2–$(SH2\text{-}B)_2$–Jak2 has been implicated in Jak2 activation (Fig. 1.6 *b*). It has been suggested that this complex might bring the two Jak2 molecules in close proximity, thereby inducing the autophosphorylation of Jak2[85, 86]. In contrast, other studies have suggested that SH2-B SH2 domain itself might activate Jak2 by inducing a conformational change of the kinase[82, 88]. Partially supported by *in vitro* experimentation, these findings mostly remain anticipatory and incomplete.

1.2.8 REFERENCES

1. Endy, D. and R. Brent, Modelling cellular behaviour. Nature, 2001. 409(6818): p. 391-395.

2. Hlavacek, W. S., J. R. Faeder, M. L. Blinov, A. S. Perelson, and B. Goldstein, The complexity of complexes in signal transduction. Biotechnology and Bioengineering, 2003. 84(7): p. 783-794.

3. Faeder, J. R., W. S. Hlavacek, I. Reischl, M. L. Blinov, H. Metzger, A. Redondo, C. Wofsy, and B. Goldstein, Investigation of Early Events in Fc{epsilon}RI-Mediated Signaling Using a Detailed Mathematical Model. Journal of Immunology, 2003. 170(7): p. 3769-3781.

4. Arkin, A.P., Synthetic cell biology. Current Opinion in Biotechnology, 2001. 12(6): p. 638-644.

5. Conzelmann, H., J. Saez-Rodriguez, T. Sauter, B. Kholodenko, and E. Gilles, A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks. BMC Bioinformatics, 2006. 7(1): p. 34.

6. Borisov, N. M., N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, Trading the micro-world of combinatorial complexity for the macro-world of protein interaction domains. Biosystems5th International Conference on Systems Biology - ICSB 2004, 2006. 83(2-3): p. 152-166.

7. Barua, D., J.R. Faeder, and J.M. Haugh, Structure-based kinetic models of modular signaling protein function: Focus on Shp2. Biophysical Journal, 2007. 92(7): p. 2290-2300.

8. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek, BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics, 2004. 20(17): p. 3289-3291.

9. Hunter, T., Signaling - 2000 and beyond. Cell, 2000. 100(1): p. 113-127.

10. Cohen, G.B., R. Ren, and D. Baltimore, Modular binding domains in signal transduction proteins. Cell, 1995. 80(2): p. 237-248.

11. Pawson, T., Protein modules and signalling networks. Nature, 1995. 373(6515): p. 573-580.

12. Bhattacharyya, R. P., A. Remenyi, B. J. Yeh, and W. A. Lim, Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. Annual Review of Biochemistry, 2006. 75 (1): p. 655-680.

13. Venter, J.C., et al., The Sequence of the Human Genome. Science, 2001. 291(5507): p. 1304-1347

14. Leo, A. and B. Schraven, Networks in signal transduction: the role of adaptor proteins in platelet activation. PlateletsPlatelets J1 - Platelets, 2000. 11(8): p. 429-445.

15. Sarmay, G., A. Angyal, A. Kertesz, M. Maus, and D. Medgyesi., The multiple function of Grb2 associated binder (Gab) adaptor/scaffolding protein in immune cell signaling. Immunology Letters Signals and Signal Processing in the Immune System, 2006. 104(1-2): p. 76-82.

16. Vondriska, T.M., J.M. Pass, and P. Ping, Scaffold proteins and assembly of multiprotein signaling complexes. Journal of Molecular and Cellular Cardiology, 2004. 37(2): p. 391-397.

17. Garrington, T.P. and G.L. Johnson, Organization and regulation of mitogen-activated protein kinase signaling pathways. Current Opinion in Cell Biology, 1999. 11(2): p. 211-218.

18. Cuevas, B. D., Y. L. Lu, M. L. Mao, J. Y. Zhang, R. LaPushin, K. Siminovitch, and G. B. Mills, Tyrosine phosphorylation of p85 relieves its inhibitory activity on phosphatidylinositol 3-kinase. Journal of Biological Chemistry, 2001. 276(29): p. 27455-27461.

19. Neel, B.G., H.H. Gu, and L. Pao, The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. Trends in Biochemical Sciences, 2003. 28(6): p. 284-293.

20. Grazioli, L., V. Germain, A. Weiss, and O. Acuto, Anti-peptide antibodies detect conformational changes of the inter-SH2 domain of ZAP-70 due to binding to the zeta chain and to intramolecular interactions. Journal of Biological Chemistry, 1998. 273(15): p. 8916-8921.

21. Yamaguchi, H. and W.A. Hendrickson, Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. Nature, 1996. 384(6608): p. 484-489.

22. Pawson, T. and P. Nash, Assembly of Cell Regulatory Systems Through Protein Interaction Domains. Science, 2003. 300(5618): p. 445-452.

23. Dueber, J. E., B. J. Yeh, K. Chak, and W. A. Lim, Reporgramming Control of an Allosteric Signaling SwitchThrough Modular Recombination. Science, 2003. 301(5641): p. 1904-1908.

24. Chan, A. C., M. Dalton, R. Johnson, G. H. Kong, T. Wang, R. Thoma, and T. Kurosaki, Activation of Zap-70 Kinase-Activity by Phosphorylation of Tyrosine-493 Is Required for Lymphocyte Antigen Receptor Function. Embo Journal, 1995. 14(11): p. 2499-2508.

25. von Willebrand, M., S. Williams, M. Saxena, J. Gilman, P. Tailor, T. Jascur, G. P. Amarante-Mendes, D. R. Green, and T. Mustelin, Modification of phosphatidylinositol 3-kinase SH2 domain binding properties by Abl- or Lck-mediated tyrosine phosphorylation at Tyr-688. Journal of Biological Chemistry, 1998. 273(7): p. 3994-4000.

26. Mcglade, C. J., C. Ellis, M. Reedijk, D. Anderson, G. Mbamalu, A. D. Reith, G. Panayotou, P. End, A. Bernstein, A. Kazlauskas, M. D. Waterfield, and T. Pawson, Sh2 Domains of the P85-Alpha Subunit of Phosphatidylinositol 3-Kinase Regulate Binding to Growth-Factor Receptors. Molecular and Cellular Biology, 1992. 12(3): p. 991-997.

27. Pei, D.H., J. Wang, and C.T. Walsh, Differential functions of the two Src homology 2 domains in protein tyrosine phosphatase SH-PTP1. Proceedings of the National Academy of Sciences of the United States of America, 1996. 93(3): p. 1141-1145.

28. Isakov, N., R. Wange, W. Burgess, J. Watts, R. Aebersold, and L. Samelson, ZAP-70 binding specificity to T cell receptor tyrosine-based activation motifs: the tandem SH2 domains of ZAP-70 bind distinct tyrosine-based activation motifs with varying affinity. Journal of Experimental Medicine, 1995. 181(1): p. 375-380.

29. Ottinger, E.A., M.C. Botfield, and S.E. Shoelson, Tandem SH2 domains confer high specificity in tyrosine kinase signaling. Journal of Biological Chemistry, 1998. 273(2): p. 729-735.

30. Panayotou, G., G. Gish, P. End, O. Troung, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, and M. D. Waterfield, Interactions between Sh2 Domains and Tyrosine-Phosphorylated Platelet-Derived Growth-Factor Beta-Receptor Sequences - Analysis of Kinetic-Parameters by a Novel Biosensors-Based Approach. Molecular and Cellular Biology, 1993. 13(6): p. 3567-3576.

31. Dueber, J.E., E.A. Mirsky, and W.A. Lim, Engineering synthetic signaling proteins with ultrasensitive input/output control. Nature Biotechnology, 2007. 25(6): p. 660-662.

32. Aris, A. and A. Villaverde, Modular protein engineering for non-viral gene therapy. Trends in Biotechnology, 2004. 22(7): p. 371-377.

33. Bu, J., A. Shaw, and A. Chan, Analysis of the Interaction of ZAP-70 and syk Protein-Tyrosine Kinases with the T-Cell Antigen Receptor by Plasmon Resonance. Proceedings of the National Academy of Sciences, 1995. 92(11): p. 5106-5110.

34. Kuhns, M.S., M.M. Davis, and K.C. Garcia, Deconstructing the Form and Function of the TCR/CD3 Complex. Immunity, 2006. 24(2): p. 133-139.

35. Weiss, A., T cell antigen receptor signal transduction: A tale of tails and cytoplasmic protein-tyrosine kinases. Cell, 1993. 73(2): p. 209-212.

36. Marie Christine Béné, What is ZAP-70? Cytometry Part B: Clinical Cytometry, 2006. 70B(4): p. 204-208.

37. Lok, L., Larry, Automatic generation of cellular reaction networks with Moleculizer 1.0. Nature biotechnology, 2005. 23(1): p. 131-136.

38. Barua, D., J.R. Faeder, and J.M. Haugh, Computational Models of Tandem Src Homology 2 Domain Interactions and Application to Phosphoinositide 3-Kinase. Journal of Biological Chemistry, 2008. 283(12): p. 7338-7345.

39. Borisov, N. M., N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, Signaling through receptors and scaffolds: Independent interactions reduce combinatorial complexity. Biophysical Journal, 2005. 89(2): p. 951-966.

40. Koschorreck, M., H. Conzelmann, S. Ebert, M. Ederer, and E. Gilles, Reduced modeling of signal transduction - a modular approach. BMC Bioinformatics, 2007. 8(1): p. 336.

41. Pawson, T. and P. Nash, Assembly of Cell Regulatory Systems Through Protein Interaction Domains. Science, 2003. 300(5618): p. 445-452.

42. Asthagiri, A.R. and D.A. Lauffenburger, A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. Biotechnology Progress, 2001. 17(2): p. 227-239.

43. Hatakeyama, M., S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J. H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama, and A. Konagaya, A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. Biochemical Journal, 2003. 373: p. 451-463.

44. Kholodenko, B. N., O. V. Demin, G. Moehren, and J. B. Hoek, Quantification of short term signaling by the epidermal growth factor receptor. Journal of Biological Chemistry, 1999. 274(42): p. 30169-30181.

45. Haugh, J. M., K. Schooler, A. Wells, H. S. Wiley, and D. A. Lauffenburger, Effect of epidermal growth factor receptor internalization on regulation of the phospholipase C-gamma 1 signaling pathway. Journal of Biological Chemistry, 1999. 274(13): p. 8958-8965.

46. Park, C.S., I.C. Schneider, and J.M. Haugh, Kinetic Analysis of Platelet-derived Growth Factor Receptor/Phosphoinositide 3-Kinase/Akt Signaling in Fibroblasts. Journal of Biological Chemistry, 2003. 278(39): p. 37064-37072.

47. Haugh, J.M., A. Wells, and D.A. Lauffenburger, Mathematical modeling of epidermal growth factor receptor signaling through the phospholipase C pathway: Mechanistic insights and predictions for molecular interventions. Biotechnology and Bioengineering, 2000. 70(2): p. 225-238.

48. Sedaghat, A.R., A. Sherman, and M.J. Quon, A mathematical model of metabolic insulin signaling pathways. American Journal of Physiology-Endocrinology and Metabolism, 2002. 283(5): p. E1084-E1101.

49. Schoeberl, B., Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nature biotechnology, 2002. 20(4): p. 370-375.

50. Bhattacharyya, R. P., A. Remenyi, B. J. Yeh, and W. A. Lim, Domains, Motifs, and Scaffolds: The Role of Modular Interactions in the Evolution and Wiring of Cell Signaling Circuits. Annual Review of Biochemistry, 2006. 75(1): p. 655-680.

51. Bradshaw, J. M., Y. Kubota, T. Meyer, and H. Schulman, An ultrasensitive Ca2+/calmodulin-dependent protein kinase II-protein phosphatase 1 switch facilitates specificity in postsynaptic calcium signaling. Proceedings of the National Academy of Sciences, 2003. 100(18): p. 10512-10517.

52. Hlavacek, W. S., J. R. Faeder, M. L. Blinov, R. G. Posner, M. Hucka, and W. Fontana, Rules for Modeling Signal-Transduction Systems. Science Signaling The Signal Transduction Knowledge Environment, 2006. 18(344): p. re6-.

53. Xiao, S., D. Rose, T. Sasaoka, H. Maegawa, T. Burke, Jr, P. Roller, S. Shoelson, and J. Olefsky, Syp (SH-PTP2) is a positive mediator of growth factor-stimulated mitogenic signal transduction. Journal of Biological Chemistry, 1994. 269(33): p. 21244-21248.

54. Zhao, Z., Z. Tan, J. H. Wright, C. D. Diltz, S.-H. Shen, E. G. Krebs, and E. H. Fischer, Altered Expression of Protein-tyrosine Phosphatase 2C in 293 Cells Affects Protein Tyrosine Phosphorylation and Mitogen-activated Protein Kinase Activation. Journal of Biological Chemistry, 1995. 270(20): p. 11765-11769.

55. Wang, N., Z. Li, R. Ding, G. D. Frank, T. Senbonmatsu, E. J. Landon, T. Inagami, and Z. J. Zhao, Antagonism or Synergism: Role of Tyrosine Phosphatases SHP-1 and SHP-2 in Growth Factor Signaling. Journal of Biological Chemistry, 2006. 281(31): p. 21878-21883.

56. Noguchi, T., T. Matozaki, K. Horita, Y. Fujioka, and M. Kasuga, Role of SH-PTP2, a protein-tyrosine phosphatase with Src homology 2 domains, in insulin-stimulated Ras activation. Molecular and Cellular Biology, 1994. 14(10): p. 6674-6682.

57. Hof, P., S. Pluskey, S. Dhe-Paganon, M. J. Eck, and S. E. Shoelson, Crystal structure of the tyrosine phosphatase SHP-2. Cell, 1998. 92(4): p. 441-450.

58. Feng, G.S., C.C. Hui, and T. Pawson, Sh2-Containing Phosphotyrosine Phosphatase as a Target of Protein-Tyrosine Kinases. Science, 1993. 259(5101): p. 1607-1611.

59. Vogel, W., et al., Activation of a Phosphotyrosine Phosphatase by Tyrosine Phosphorylation. Science, 1993. 259(5101): p. 1611-1614.

60. Bennett, A.M., et al., Protein-Tyrosine-Phosphatase Shptp2 Couples Platelet-Derived Growth-Factor Receptor-Beta to Ras. Proceedings of the National Academy of Sciences of the United States of America, 1994. 91(15): p. 7335-7339.

61. Lechleider, R. J., S. Sugimoto, A. M. Bennett, A. S. Kashishian, J. A. Cooper, S. E. Shoelson, C. T. Walsh, and B. G. Neel, Activation of the Sh2-Containing Phosphotyrosine Phosphatase Sh-Ptp2 by Its Binding-Site, Phosphotyrosine-1009, on the Human Platelet-Derived Growth-Factor Receptor-Beta. Journal of Biological Chemistry, 1993. 268(29): p. 21478-21481.

62. Dechert, U., M. Adam, K. W. Harder, I. Clarklewis, and F. Jirik, Characterization of Protein-Tyrosine-Phosphatase Sh-Ptp2 - Study of Phosphopeptide Substrates and Possible Regulatory Role of Sh2 Domains. Journal of Biological Chemistry, 1994. 269(8): p. 5602-5611.

63. Sugimoto, S., T. J. Wandless, S. E. Shoelson, B. G. Neel, and C. T. Walsh, Activation of the Sh2-Containing Protein-Tyrosine-Phosphatase, Sh-Ptp2, by Phosphotyrosine-Containing Peptides Derived from Insulin-Receptor Substrate-1. Journal of Biological Chemistry, 1994. 269(18): p. 13614-13622.

64. Lu, W., D. Q. Gong, D. Bar-Sagi, and P. A. Cole, Site-specific incorporation of a phosphotyrosine mimetic reveals a role for tyrosine phosphorylation of SHP-2 in cell signaling. Molecular Cell, 2001. 8(4): p. 759-769.

65. Lu, W., K. Shen, and P.A. Cole, Chemical dissection of the effects of tyrosine phosphorylation of SHP-2. Biochemistry, 2003. 42(18): p. 5461-5468.

66. Zhang, Z. S., K. Shen, W. Lu, and P. A. Cole, The role of C-terminal tyrosine phosphorylation in the regulation of SHP-1 explored via expressed protein ligation. Journal of Biological Chemistry, 2003. 278(7): p. 4668-4674.

67. Kazlauskas, A., The 64-Kda Protein That Associates with the Platelet-Derived Growth-Factor Receptor Beta-Subunit Via Tyr-1009 Is the Sh2-Containing Phosphotyrosine Phosphatase Syp. Proceedings of the National Academy of Sciences of the United States of America, 1993. 90(15): p. 6939-6942.

68. Klinghoffer, R.A. and A. Kazlauskas, Identification of a Putative Syp Substrate, the Pdgf-Beta Receptor. Journal of Biological Chemistry, 1995. 270(38): p. 22208-22217.

69. Ekman, S., A Kallin, A Engström, and C-H Heldin, SHP-2 is involved in heterodimer specific loss of phosphorylation of Tyr771 in the PDGF beta-receptor. Oncogene, 2002. 21(12): p. 1870-1875.

70. Zhao, R. and Z.J. Zhao, Tyrosine phosphatase SHP-2 dephosphorylates the platelet-derived growth factor receptor but enhances its downstream signalling. Biochemical Journal, 1999 338(1): p. 35-39.

71. Leevers, S.J., B. Vanhaesebroeck, and M.D. Waterfield, Signalling through phosphoinositide 3-kinases: the lipids take centre stage. Current Opinion in Cell Biology, 1999. 11(2): p. 219-225.

72. Vanhaesebroeck, B. and M.D. Waterfield, Signaling by Distinct Classes of Phosphoinositide 3-Kinases. Experimental Cell Research, 1999. 253(1): p. 239-254.

73. Cantley, C., D. Fruman, C. Yballe, and L. Rameh, Role of PI 3-kinase in cell signaling. Faseb Journal, 1999. 13(7): p. A1583-A1583.

74. Rameh, L.E. and L.C. Cantley, The role of phosphoinositide 3-kinase lipid products in cell function. Journal of Biological Chemistry, 1999. 274(13): p. 8347-8350.

75. Shibasaki, F., Y. Homma, and T. Takenawa, 2 Types of Phosphatidylinositol 3-Kinase from Bovine Thymus - Monomer and Heterodimer Form. Journal of Biological Chemistry, 1991. 266(13): p. 8108-8114.

76. Carpenter, C.L., et al., Purification and Characterization of Phosphoinositide 3-Kinase from Rat-Liver. Journal of Biological Chemistry, 1990. 265(32): p. 19704-19711.

77. Panayotou, G., et al., Interaction of the P85-Subunit of Pi 3-Kinase and Its N-Terminal Sh2-Domain with a Pdgf Receptor Phosphorylation Site - Structural Features and Analysis of Conformational-Changes. Embo Journal, 1992. 11(12): p. 4261-4272.

78. Carpenter, C. L., B. C. Duckworth, K. R. Auger, B. Cohen, B. S. Schaffhausen, and L. C. Cantley, Phosphoinositide 3-Kinase Is Activated by Phosphopeptides That Bind to the Sh2 Domains of the 85-Kda Subunit. Journal of Biological Chemistry, 1993. 268(13): p. 9478-9483.

79. Shoelson, S. E., M. Sivaraja, K. P. Williams, P. Hu, J. Schlessinger, and M. A. Weiss, Specific Phosphopeptide Binding Regulates a Conformational Change in the Pi 3-Kinase Sh2 Domain Associated with Enzyme Activation. Embo Journal, 1993. 12(2): p. 795-802.

80. Kavanaugh, W. M., A. Klippel, J. A. Escobedo, and L. T. Williams, Modification of the 85-Kilodalton Subunit of Phosphatidylinositol-3 Kinase in Platelet-Derived Growth Factor-Stimulated Cells. Molecular and Cellular Biology, 1992. 12(8): p. 3415-3424.

81. Piccione, E., R. D. Case, S. M. Domchek, P. Hu, M. Chaudhuri, J. M. Backer, J. Schlessinger, and S. E. Shoelson, Phosphatidylinositol 3-Kinase P85 Sh2 Domain Specificity Defined by Direct Phosphopeptide Sh2 Domain Binding. Biochemistry, 1993. 32(13): p. 3197-3202.

82. Kurzer, J. H., P. Saharinen, O. Silvennoinen, and C. Carter-Su, Binding of SH2-B Family Members within a Potential Negative Regulatory Region Maintains JAK2 in an Active State. Molecular and Cellular Biology, 2006. 26(17): p. 6381-6394.

83. Kurzer, J. H., L. S. Argetsinger, Y.-J. Zhou, J.-L. K. Kouadio, J. J. O'Shea, and C. Carter-Su, Tyrosine 813 Is a Site of JAK2 Autophosphorylation Critical for Activation of JAK2 by SH2-B{beta}. Molecular and Cellular Biology, 2004. 24(10): p. 4557-4570.

84. Rui, L. and C. Carter-Su, Identification of SH2-Bbeta as a potent cytoplasmic activator of the tyrosine kinase Janus kinase 2. Proceedings of the National Academy of Sciences, 1999. 96(13): p. 7172-7177.

85. Nishi, M., et al., Kinase Activation through Dimerization by Human SH2-B. Molecular and Cellular Biology, 2005. 25(7): p. 2607-2621.

86. Dhe-Paganon, S., E. D. Werner, M. Nishi, L. Hansen, Y.-I. Chi, and S. E. Shoelson, A phenylalanine zipper mediates APS dimerization. Nature Structural & Molecular Biology, 2004. 11(10): p. 968-974.

87. Rui, L., J. Herrington, and C. Carter-Su, SH2-B, a Membrane-associated Adapter, Is Phosphorylated on Multiple Serines/Threonines in Response to Nerve Growth Factor by Kinases within the MEK/ERK Cascade, Journal of Biological Chemistry, 1999. 274(37): p. 26485-26492.

88. Rui, L., D. R. Gunter, J. Herrington, and C. Carter-Su, Differential Binding to and Regulation of JAK2 by the SH2 Domain and N-Terminal Region of SH2-Bbeta. Molecular and Cellular Biology, 2000. 20(9): p. 3168-3177.

CHAPTER 2

BioNetGen2: software for rule-based computational modeling

2.1 BioNetGen AND ITS APPROACH

BioNetGen is a software package for the modeling and simulation of biochemical networks in which proteins (and other biomolecules) are represented as structured objects and interactions among these objects are described by rules. The motivation for this modeling language is that a small number of proteins, each comprised of a few functional domains and interacting in a small number of different ways, can give rise to very large networks of chemical species – each representing a distinct protein complex – and reactions – each representing a distinct transformation among species. As all the models presented in this study illustrate, this phenomenon, which has been called combinatorial complexity [1], is endemic to cell signaling biochemistry. The conventional approach of writing the network equations by hand (or using one of the many available network drawing tools) is untenable without making a number of simplifying assumptions, which are usually not justified by

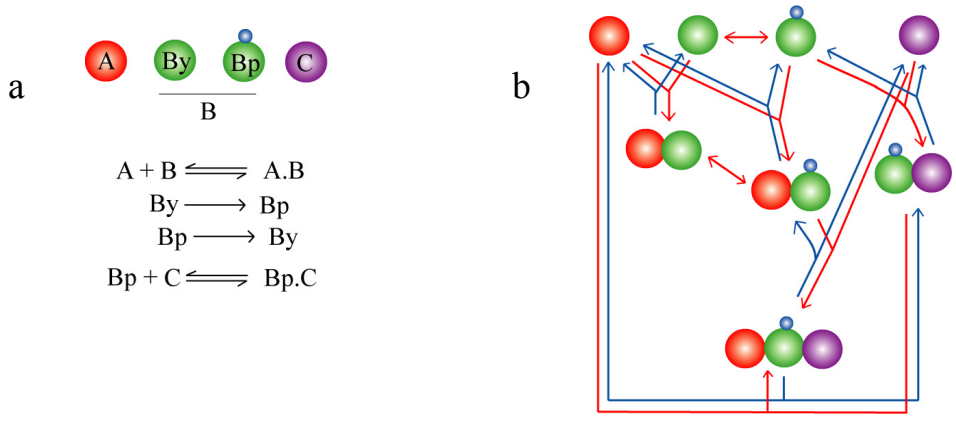


FIGURE 2 **Rule-based scheme of network generation.** (*a*) A hypothetical model system is shown with three 'seed' molecules and their allowable states and reactions. A rule-based software requires these molecules and reactions to be written in a machine-readable language. Each reaction shown represents a separate rule specifying a generic interaction/transformation of species. (*b*) Combinatorial species and their connectivities generated upon execution of the generic reaction rules are shown.

experimental information and which can lead to unanticipated effects on network structure [2]. The alternative approach, which has been pursued recently by a number of groups (reviewed in [1]), uses a relatively small set of rules that describe the protein-protein interactions present in the system. These are processed by a computer program to generate the full network of species and reactions (Fig. 2). If the language used to express the rules is well constructed, they provide a concise description of the signaling network that is easily extended to encompass new elements – including new proteins, new components of existing proteins, and new protein-protein interactions. The major drawbacks of this approach are that new software tools are required to interpret the language and potentially very large networks of possible species and reactions must be simulated. A number of software tools for rule-based modeling are currently available, including StochSim [3], BioNetGen [4], Moleculizer [5], BIOCHAM [6], and Simmune [7], and these are being actively developed to address such issues as ease-of-use, simulation efficiency, and simulation of spatial effects.

In this study we have used version 2 of the BioNetGen to develop rule-based models of Shp2 regulation. This software and supporting documentation are freely available and may be downloaded from http://bionetgen.lanl.gov. BioNetGen2 (BNG2) uses a graphical notation to represent proteins and complexes and graph-rewriting rules to describe protein-protein interactions and other biochemical transformations [8]. The theoretical basis for this representation and algorithms used in the implementation of the software have been described previously [9]. Here, in this chapter, we provide a brief description of the rule-building syntax and conventions of BNG2 language in terms of two models of the Shp2 regulation discussed in chapter three and presented in the input files, Shp2_base.bngl and Shp2_extended.bngl in Appendix A.

2.2 SYNTAX AND GRAMMAR OF BioNetGen LANGUAGE

2.2.1 Representation of molecules and complexes

In the BNG2 language proteins and other biomolecules are represented as structured objects called Molecules. Each Molecule may contain any number of Components that represent structural or functional elements of the protein, such as protein domains and

phosphorylation sites. Components may have internal states that may, for example, represent a conformational state of a domain or a phosphorylation state. The Shp2 phosphatase in the base model is declared in the following way

```
S(NSH2~C~O,CSH2,PTP~C~O)
```

which means that Molecules of type S (for Shp2) are comprised of three Components – NSH2, CSH2, and PTP. The NSH2 and PTP Components may each be present in either a closed (C) or open (O) state, whereas the CSH2 Component does not have an associated state variable. It is important to emphasize that only internal states of Components are declared; binding states of Components arise from interactions that are specified in species declarations and reaction rules, as described below. The receptor in both models is declared as

```
R(DD,Y1~U~P,Y2~P)
```

which indicates that the Y1 component has both an unphosphorylated (U) state and a phosphorylated (P) state. The Y2 Component, on the other hand, is only found in the phosphorylated state.

Molecules may bind to other Molecules to form complexes. Bonds are indicated as links that connect two Components in the associating Molecules. Bonds are usually created through the application of reaction rules, but it is also sometimes useful to specify complexes with bonds that are unbreakable for a particular rule set. The current models take the number of dimeric receptor complexes to be an input variable, and thus it is useful to define this complex in the input file as

```
R(DD!1,Y1~U,Y2~P).R(DD!1,Y1~U,Y2~P)
```

where the two receptors in the complex are bound through the link between the dimerization domain (DD) of each receptor. The '.' is used to group molecules into a complex; however, the association is invalid if there is no bond connecting at least one component of each Molecule to the rest of the complex. The two Components linked through a bond are indicated by an '!' followed by the same arbitrary index. Here, a bond with index 1 links the DD components of the two receptors. The scope of bond indices is local to the complex in which they are used. Bonds between components in the same molecule are also allowed.

2.2.2 Input file structure and rule-writing convention

A BNG2 model is comprised of the following input blocks, each of which begins with the line `begin` *`blockname`* and ends with the line `end` *`blockname`*: `parameters`, `molecule types`, `seed species`, `reaction rules`, and `observables`.

The `parameters` block is used for the declaration of numerical parameters that are used to designate the initial concentrations of species and rates constants used in reaction rules. Each parameter is declared on a separate line with the syntax *parameter_name value*.

The `molecule types` block is used for the declaration of Molecules, as discussed above. This block is optional, but highly recommended because it allows more comprehensive error checking that reduces the likelihood of unintended consequences from the model specification. Each Molecule is declared on a separate line.

The `seed species` block is used for the declaration of the chemical species that are used to seed network generation, which is performed by application of the reaction rules to the seed set of species followed by iterative application of the rules to the species generated by the previous iteration. A minimal set of seed species is the set of Molecules defined in the model. Note that any Component that has an associated state variable must be in a defined state. Each species is declared on a separate line followed by its initial concentration, which may be either a parameter name or a numerical value.

The `reaction rules` block is the heart of the BNG2 model and is used to define the biochemical events that can take place. Each Rule is declared on a separate line. Two basic types of Rules are illustrated in the example files provided: Rules that make or break a bond between two components and Rules that change the internal states of one or more components. An example of the latter is the Rule that generates intra-complex phosphorylation

```
R(DD!+,Y1~U) -> R(DD!+,Y1~P)  kkin_Y1
```

which illustrates a number of important elements of Rule syntax. In words, this Rule states that a Molecule R that has a bound DD Component and a Y1 Component (shown in red) that is unbound and in its U state undergoes transformation of Y1 to its P state with first order kinetics given by the rate constant kkin_Y1. The '!+' following the DD Component indicates

that this Component must participate in at least one bond. Any Component that is listed in a Rule is assumed to be unbound unless another binding state is indicated using either an explicit bond or a bond wildcard. The other bond wildcard is '!?', which indicates that a Component may be bound but is not required to be. The unidirectional arrow indicates that the Rule is applied only in the forward direction. A key aspect of this Rule is the omission of the Y2 Component of R, which means that the Rule is applied regardless of the internal or binding state of the Y2 Component. In the Shp2 regulation models, this Rule embodies the assumption that the binding of Shp2 to Y2 has no effect on the intrinsic activity of the RTK to phosphorylate Y1.

The input files contain Rules for two basic types of binding reactions – intermolecular and intramolecular. These types can be distinguished by the number of reactants involved. For an intermolecular reaction, patterns selecting the reactants are separated by a '+' in the standard way for a chemical reaction, as in the rule

```
R(Y2~P) + S(CSH2) <-> R(Y2~P!1).S(CSH2!1)  kon_CSH2,koff_CSH2 \
  exclude_reactants(2,R)
```

which describes binding of the Shp2 C-SH2 domain to receptors phosphorylated on Y2. The presence of the '+' operator here requires that the bond form between an R and an S from two separate complexes. If this were an intramolecular reaction, the '.' operator would be used instead, as illustrated below. This rule also illustrates the syntax of a reversible reaction, in which reactants and products are linked by a bi-directional arrow. Two rate constants, for the forward and reverse directions respectively, are provided. The rule also has an additional modifier on the second line of the rule, which is extended by the continuation character, '\', placed at the end of the first line. The directive

`exclude_reactants(`*`reactant_index,reactant_pattern1,…`*`)`

excludes species matching the specified pattern(s) from participating in the reaction as the reactant specified by the index. In the Rule given above the second reactant pattern is `S(CSH2)`, which selects Shp2 molecules with an unbound CSH2 component. The `exclude_reactants` directive prevents species that contain and Shp2 with a free CSH2 but also contain at least one R Molecule from being acted on by the Rule. This restriction is used to ensure that the Shp2 undergoing the reaction is cytosolic and not membrane-

associated, which is indicated by the presence of R in the complex. This also prevents Shp2 that is already bound to a receptor in a complex from cross-linking receptor complexes.

There are several points to consider when this Rule is applied in the reverse direction to break a bond between R and S. First, just as the '+' operator required that R and S be part of two separate species for bond formation, it requires that the reverse reaction form two separate species when the bond is broken. Thus, if breaking the specified bond between a particular R and S in a complex does not break the complex into two parts, the Rule is not applied. A second Rule would be required for the case in which breaking the bond does not separate R and S into distinct complexes. The second point is that the `exclude_reactants` directive is automatically converted to an `exclude_products` directive when the Rule is applied in the reverse direction. Thus, the Rule would not be applied if breaking a particular bond between R and S left S in a complex with a different R molecule.

The great majority of the Rules for the Shp2 regulation models (16 out of 22 Rules for the base model) involve intramolecular binding events. Although it is possible to represent the network with as few as three intramolecular Rules, the use of a larger number of more complex Rules permits the intracomplex binding factors, χ_r, to depend on the geometry of the complex. A typical intracomplex binding Rule is

```
R(Y1~P,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O) <-> \
R(Y1~P!2,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O!2) \
chi_r5*kon_PTP,koff_PTP
```

which describes the association of a phosphorylated Y1 component of an R Molecule with a PTP domain of an S Molecule when the Y2 Component of the same R is phosphorylated and bound to the NH2 Component (open configuration) of the same S Molecule. Because R and S are linked by the '.' operator, the Rule requires that R and S also be in the same complex, both when the bond is formed and when it is broken. For clarity, the Components directly affected by the reaction are highlighted in red. This Rule also illustrates the use of a multiplicative factor, chi_r5, to modify the value of the rate constant.

The flexibility of the BioNetGen modeling language is demonstrated by considering the steps required to extend the base model (Shp2_base.bngl) to include additional regulation of Shp2 through phosphorylation of a C-terminal tyrosine residue (Shp2_extended.bngl). The

extension requires three additional Reaction Rules, their associated rate parameters, and modification of the Molecule Types and Seed Species declarations for S. The declaration of S is changed to

S(NSH2~C~O,CSH2,PTP~C~O,Y~U)

where the additional Component Y (shown in red) represents a C-terminal residue or residues. The additional Rules describe the phosphorylation and dephosphorylation of the Y Component of Shp2 and its intramolecular binding in the phosphorylated state to the open conformation of the NSH2 domain. The modest addition of these model elements increases the size of the generated network from 149 species to 1325 species.

The `observables` block is used for the declaration of variables that are defined as sums of the concentrations of species that match a pattern or a set of patterns and which are useful for defining outputs of a model. The single Observable for both models considered here is defined by

```
Molecules  pYR      R(Y1~P!?)
```

which gives the total level of phosphorylation of the Y1 component of R in the system. The binding state wildcard '!?' indicates that the binding state of the Y1 Component does not affect the match (in contrast to the pattern `R(Y1~P)`, which would select only phosphorylated Y1 Components that are not bound). The Molecules keyword here indicates that the Observable is of the type Molecules, which means that the concentration of each species selected for the sum is multiplied by the number of occurrences of the pattern in that species. Thus, the concentration of a dimer with two Molecules of R phosphorylated at Y1 would be weighted by a factor of two. Observables of type Species produce an unweighted sum over the matching species.

Following the specification of the model, commands can be issued to generate the reaction network and perform simulations based on deterministic (ordinary differential equations (ODEs)) or stochastic (Gillespie's direct method [10]) algorithms. Only ODE-based simulations are used in the current work. The sequence of commands used for both models is

```
generate_network();
writeSBML();
simulate_ode({t_end=>1000,n_steps=>100,steady_state=>1,atol=>1e-10,rtol=>1e-12});
```

where `generate_network` applies the Reaction Rules to the seed set of species in the iterative manner outlined above, `writeSBML` generates a model file in Systems Biology Markup Language (SBML) level 2 format [11] for export to other applications, and `simulate_ode` computes the time course of species concentrations and Observables from time t = 0 to time t = `t_end` sampling `n_steps` times. Setting the `steady_state` parameter to a non-zero value indicates that the simulation will be stopped before `t_end` is reached if the concentrations satisfy a convergence criterion, which is based on the absolute integration tolerance `atol`. `rtol` is the relative tolerance used in the integration of the ODEs. Additional parameters controlling the behavior of each command as well as the syntax of other commands can be found in the online documentation at http://bionetgen.lanl.gov.

Running the `BNG2.pl` command on the `SHP_base.bngl` generates descriptive output in the terminal window (stdout and stderr in Unix-like environments) and the files `SHP_base.net`, `SHP_base.xml`, `SHP_base.cdat`, and `SHP_base.gdat`. The NET file contains the generated network in BNG2 format, the XML file contains the same network in SBML format, the CDAT file contains the time courses of all species in tabular format, and the GDAT file contains the time courses of the observables in tabular format. A utility called `PhiBPlot.jar` is included in the BioNetGen distribution to facilitate plotting of the CDAT and GDAT files.

2.3 REFERENCES

1. Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, et al. (2006) Rules for modeling signal-transduction systems. Science STKE 2006: re6.

2. Blinov ML, Faeder JR, Goldstein B, Hlavacek WS (2006) A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. Biosystems 83: 136-151.

3. Le Novere N, Shimizu TS (2001) STOCHSIM: Modelling of stochastic biomolecular processes. Bioinformatics 17: 575-576.

4. Blinov ML, Faeder JR, Goldstein B, Hlavacek WS (2004) BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics 20: 3289-3291.

5. Lok L, Brent R (2005) Automatic generation of cellular networks with Moleculizer 1.0. Nat Biotechnol 23: 131-136.

6. Calzone L, Fages F, Soliman S (2006) BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge. Bioinformatics 22: 1805-1807

7. Meier-Schellersheim M, Xu X, Angermann BR, Kunkel EJ, Jin T, et al. (2006) Key role of local regulation in chemosensing revealed by a new molecular interaction-based modeling method. PLoS Computat Biol: e82

8. Faeder JR, Blinov ML, Hlavacek WS (2005) Graphical rule-based representation of signal-transduction networks. Proc ACM Symp Appl Computing: 133-140.

9. Blinov ML, Yang J, Faeder JR, Hlavacek WS (2005) Graph theory for rule-based modeling of biochemical networks. Proc BioCONCUR 2005.

10. Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comp Phys 22: 403-434.

11. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19: 524-531.

CHAPTER 3

Structure-based kinetic models of modular signaling protein function: focus on Shp2

Adapted from Barua, Faeder and Haugh, Biophys. J., 92: 2290-2300 (2007)

3.1 ABSTRACT

We present here a computational, rule-based model to study the function of the SH2 domain-containing protein-tyrosine phosphatase, Shp2, in intracellular signal transduction. The two SH2 domains of Shp2 differentially regulate the enzymatic activity by a well-characterized mechanism, but they also affect the targeting of Shp2 to signaling receptors in cells. Our kinetic model integrates these potentially competing effects by considering the intra- and intermolecular interactions of the Shp2 SH2 domains and catalytic site as well as the effect of Shp2 phosphorylation. Even for the isolated Shp2/receptor system, which may seem simple by certain standards, we find that the network of possible binding and phosphorylation states is comprised of over one thousand members. To our knowledge, this is the first kinetic model to fully consider the modular, multifunctional structure of a signaling protein, and the computational approach should be generally applicable to other complex intermolecular interactions.

3.2 INTRODUCTION

Intracellular signal transduction pathways mediated by growth factor, cytokine, and hormone receptors control cell behavior, such as proliferation, survival, migration, and differentiation. Protein tyrosine kinases, including growth factor receptors of the receptor tyrosine kinase (RTK) class, catalyze tyrosine phosphorylation and are typically involved in the initial steps of signal transduction, but equally important in determining the magnitude and kinetics of the intracellular response are the opposing tyrosine dephosphorylation reactions, catalyzed by protein tyrosine phosphatases (PTPs) [1]. In RTK signaling, the ligated receptors dimerize such that their intracellular tails may be trans-phosphorylated on specific tyrosine residues, which then act as a scaffold for the binding of cytosolic signaling proteins with Src homology 2 (SH2) and other phosphotyrosine-binding domains. These proteins include certain enzymes, adaptor proteins, and transcription factors, which may be activated upon binding to the receptor complex and/or by subsequent phosphorylation, dephosphorylation, or other modifications [2, 3]. In general, the modular design of signaling proteins, typically comprised of multiple domains with specific functions, is a central paradigm in the understanding of signal transduction networks; how these domains work together in the full-length protein, and in the context of signaling mediated by certain receptors, is a problem steeped in complexity [4-6].

In this paper, we analyze the function of one such signaling protein, SH2 domain-containing phosphatase (Shp) 2. Shp2 and the closely related Shp1 are intriguing because each has a classic PTP catalytic domain as well as two SH2 domains that regulate its activity and help target the enzyme to phosphorylated RTKs and other tyrosine-phosphorylated proteins [7]. Whereas most PTPs are thought to simply antagonize RTK signaling, Shp2 plays a positive role in the activation of the Ras/extracellular signal-regulated kinase (Erk) signaling cascade mediated by platelet-derived growth factor (PDGF) receptors, epidermal growth factor (EGF) receptor, and other RTKs. A candidate mechanism for this effect is the dephosphorylation of the tyrosine residues in PDGF and EGF receptors that bind Ras GTPase-accelerating protein (RasGAP), a negative regulator of Ras activation; these residues

are distinct from those that engage the SH2 domains of Shp2 [8-10]. The precise dephosphorylation targets and signaling roles of Shp2 in cells have yet to be fully elucidated, however.

The structure of Shp2 and its relationship to the regulation of catalytic activity, at least in solution, are relatively well understood. The crystal structure of Shp2 clearly shows an auto-inhibition of the PTP catalytic site by the N-terminal SH2 domain (N-SH2), while the more C-terminal SH2 domain (C-SH2) does not interact in this way [11, 12]. The auto-inhibited, "closed" conformation is highly favored under basal conditions. Binding of the N-SH2 domain to a phosphotyrosine-containing protein stabilizes the "open" conformation of Shp2, activating the enzyme. In solution, addition of small, phosphorylated peptide sequences mimicking SH2 binding sites leads to a > 10-fold increase in Shp2 enzyme activity, and a variant with the N-SH2 domain deleted exhibits an even higher activity that is not augmented further by phosphopeptide addition [13-16]. A notable caveat in such studies is that, at least in certain cases, N-SH2-binding peptides can also be dephosphorylated by the PTP active site, requiring a more sophisticated biochemical analysis [17]. Consistent with the crystal structure, deletion of the C-SH2 domain does not activate the enzyme, nor does it influence enzyme activation by N-SH2-binding peptides in solution; however, C-SH2 deletion does reduce the potency of synthetic peptides bearing two SH2 domain-binding phosphotyrosines, suggesting that the N-SH2 and C-SH2 domains bind such peptides and multiply phosphorylated protein complexes in a synergistic manner to better stabilize the open form of Shp2 [16, 18-21].

Shp2 is also subject to tyrosine phosphorylation on two sites near its C-terminus in response to growth factor stimulation [22, 23]. Mutating these sites to phenylalanine leads to a modest reduction in Erk signaling [24], suggesting that Shp2 phosphorylation plays a role in enzyme activation. Selective attachment of chemical moieties that mimic the effects of phosphate addition yields higher Shp2 activity in solution and when microinjected into cells [25-27], and a conceptual model has emerged in which the N-SH2 domain, alone or in tandem with the C-SH2 domain, are engaged by the C-terminal phosphorylation sites in the same molecule. This hypothesis remains controversial because it cannot be directly tested

(Shp2 dephosphorylates itself in solution) and because the phosphorylation site(s) can serve an adaptor function that might impact Ras activation [28, 29].

The complexity of signal transduction has motivated the use of quantitative, mathematical modeling approaches in recent years to better understand the kinetic mechanisms involved and how they might work in concert inside the cell. Modeling can be a powerful tool for analysis, but an appropriate balance must always be struck between the inclusion of known signaling interactions and the tractability of the model [30, 31]. As one adds protein binding and phosphorylation states to the model, the number of species increases dramatically, an issue termed *combinatorial complexity* [32]. Even in the early stages of intracellular signaling, in which the formation of receptor complexes may be governed by just a handful of simple binding rules (e.g., a cytosolic protein may associate with a receptor if a specific receptor site is phosphorylated and unoccupied), the number of different combinations of protein states can easily reach the hundreds or thousands, as seen in the recent model of FcεRI signaling [33]. Computational tools for rule-based modeling, which generate the network of rate processes and construct the governing equations automatically, are becoming increasingly available [34-37].

In this work, we have constructed a rule-based kinetic model of the interactions between Shp2 and a dimerized RTK, although the analysis may be generalized to the interactions of Shp2 or Shp1 with other multi-protein complexes. Among other effects, we have systematically analyzed the dual role of the Shp2 SH2 domains; on the one hand, they regulate the enzymatic activity, as characterized in solution, while on the other hand, they target the enzyme to receptor complexes in cells. The latter influences the activity towards substrates in the complex through an induced proximity effect, as considered in our previous model of PTP regulation [38]. Accordingly, in the context of the currently accepted mechanisms of Shp2 regulation, we find regimes of receptor-Shp2 binding where N-SH2 deletion or C-terminal phosphorylation of Shp2 would either diminish or enhance receptor dephosphorylation, while deletion of the C-SH2 domain or both SH2 domains unilaterally impairs this function.

3.3 METHODS

3.3.1 Structure-based kinetic model of Shp2 regulation and function

In our models, activated receptors and Shp2 molecules are present at total concentrations of R_{tot} and S_{tot}, respectively, calculated on the basis of the cytosol volume. For the sake of simplicity, we introduce activated receptors as pre-formed dimers (dimer concentration is $R_{tot}/2$) and do not include the processes of external receptor-ligand binding and dimer association/dissociation. We are concerned only with the steady state behavior, and a dose response curve is implied by performing model calculations with different values of R_{tot} [38]. The hypothetical receptor has two phosphorylation sites: Y_1 is a substrate of the Shp2 catalytic (PTP) domain, while Y_2 engages SH2 domains of Shp2 and is always phosphorylated in our model. In the human PDGF β-receptor, Y_1 and Y_2 would correspond to Tyr 771 and Tyr 1009, for example [8, 9, 13, 39]. The assumed binding/modification rules are illustrated in Fig. 3.1 and are summarized as follows; base-case rate constant values are given in Table 3.1.

Shp2 is auto-inhibited through the reversible association of its N-SH2 and PTP domains (the closed conformation), provided that neither domain is receptor-bound; from thermodynamic constraints, it follows that neither the N-SH2 nor PTP domain can engage receptor sites when Shp2 is in the closed conformation [11], whereas it is assumed that binding of the C-SH2 domain to receptors is not affected by the transition between open and closed conformations (Fig. 3.1 *a*). Phosphorylation of receptor Y_1 sites by the intrinsic kinase activity is modeled as a single, unimolecular step, as considered in previous models [33, 40], whereas dephosphorylation of Y_1 by Shp2 is modeled with explicit accounting of the enzyme-substrate interaction and catalytic step (Fig. 3.1 *b*).

Once an Shp2 molecule is recruited from the cytosol to an activated receptor complex, its free SH2 and/or PTP domains may associate with free phosphotyrosine sites in the complex. As has been suggested in the literature for PDGF β-receptor [20], we allow the two SH2 domains of Shp2 to bridge the two Y_2 sites in the receptor dimer. Such ring closure transitions are unimolecular, and their rate constants are calculated by multiplying the second-order rate constants characterizing the corresponding bimolecular associations by

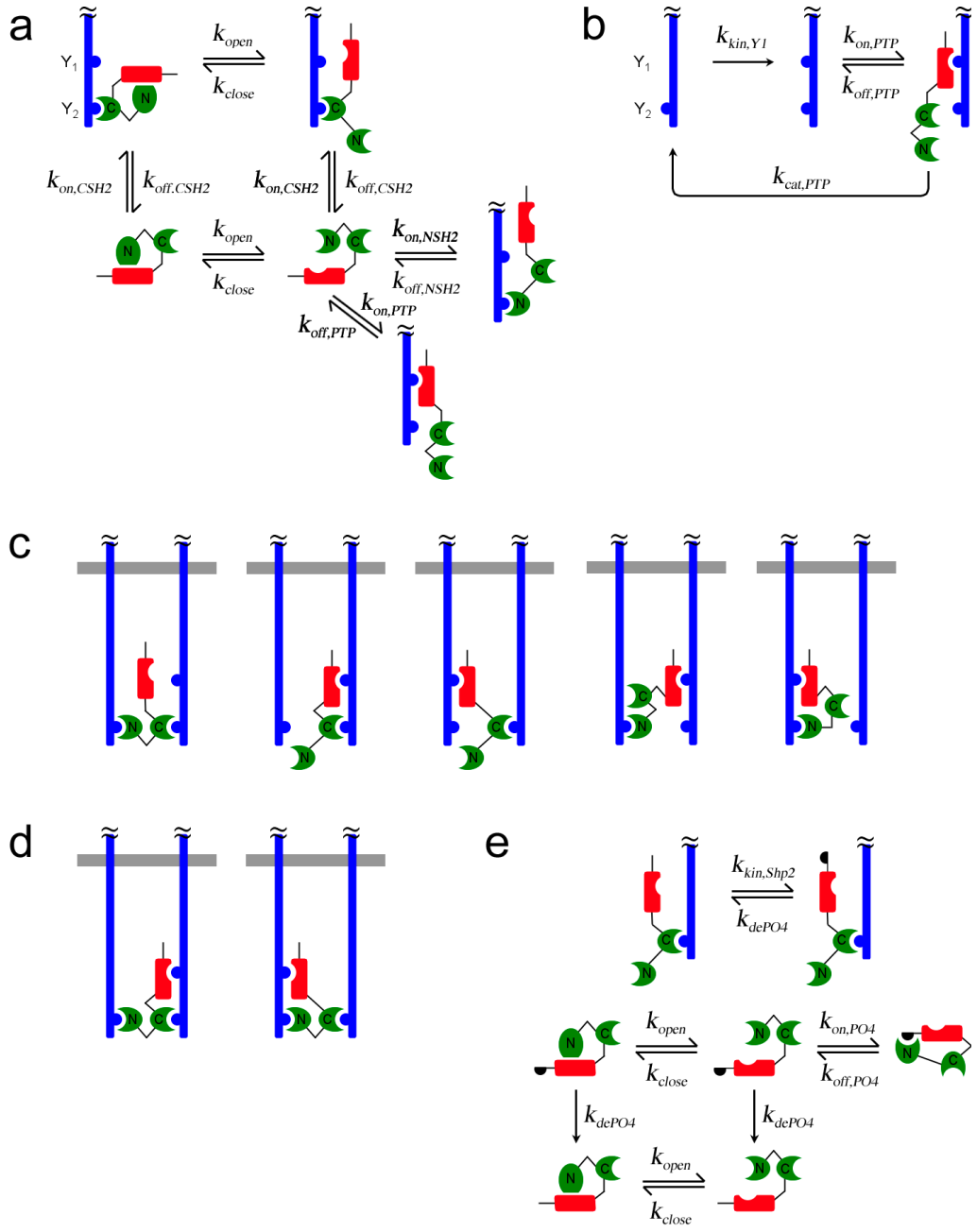


FIGURE 3.1 Kinetic model of Shp2 structure-function in receptor signaling. The hypothetical signaling complex contains two activated, dimerized receptors (blue), each with two tyrosine phosphorylation sites considered: phosphorylated Y_1 is a substrate of the Shp2 catalytic PTP domain (red), while phosphorylated Y_2 engages either SH2 domain (green) of Shp2. (*a*) Recruitment of Shp2 from the cytosol. The PTP and N-SH2 domains may only bind receptors when they are in the "open" conformation. (*b*) Phosphorylation and Shp2-mediated dephosphorylation of the receptor Y_1 site. (*c*) The five types of complexes with two Shp2 domains engaged by receptor sites. (*d*) The two types of complexes with three Shp2 domains engaged by receptor sites. (*e*) Receptor-mediated phosphorylation of Shp2 favors the open PTP conformation by occupying the N-SH2 domain.

Table 3.1 Base-case values of rate constants used in model calculations

Rate constant	Base value(s)
k_{open}	10 s^{-1}
k_{close}	500 s^{-1}
$k_{on,CSH2} = k_{on,NSH2}$	1 $\mu M^{-1}s^{-1}$ [58]
$k_{off,CSH2} = k_{off,NSH2}$	0.1-10 s^{-1} [20, 42]
$k_{on,PTP}$	1 $\mu M^{-1}s^{-1}$ [58]
$k_{off,PTP}$	10 s^{-1} [14, 59]
$k_{cat,PTP}$	1 s^{-1} [14, 59]
$k_{kin,Y1}$	0.1 s^{-1}
$k_{kin,Shp2}$ †	1 s^{-1}
k_{dePO4} †	1 s^{-1}
$k_{on,PO4}$ †	100 s^{-1}
$k_{off,PO4}$ †	0.1 s^{-1}

† Only applicable to the extended model, presented in Fig. 3.5.

various conversion factors χ_r, which have units of concentration and ensure that microscopic reversibility is satisfied [38]. If one were to assume that binding partners in the same complex are confined within a volume of 100 nm^3 (10^{-20} L), the concentration of a single molecule in that volume is ~ 10 mM; this is a reasonable estimate of χ_r, although its value might be significantly increased or decreased subject to orientation constraints. The present model allows different values of χ_r for different interactions but also recognizes that not all of these values are independent. There are 5 types of complexes with two Shp2 domains bound (Fig. 3.1 *c*) and 2 types of complexes with all 3 Shp2 domains occupied (Fig. 3.1 *d*); the overall equilibrium constant for the formation of a complex requiring multiple steps does not depend on the order of the steps, and 7 independent χ_r factors were identified and assigned conservative base-case values of 0.1 or 1 mM (Table 3.2, and Fig. A4, Appendix A). One might take as a base assumption that all such factors have the same value, as

Table 3.2 Conversion factors for intra-complex binding of Shp2 domains

Order of binding	χ_r factors (value in μM)
C-Y_2; *N*-Y_2'; *P*-Y_1	χ_{r1} (10^3); χ_{r6} (10^2)
C-Y_2; *N*-Y_2'; *P*-Y_1'	χ_{r1} (10^3); χ_{r7} (10^2)
C-Y_2; *P*-Y_1; *N*-Y_2'	χ_{r2} (10^2); $\chi_{r8} = \chi_{r1}\chi_{r6}/\chi_{r2}$
C-Y_2; *P*-Y_1'; *N*-Y_2'	χ_{r3} (10^3); $\chi_{r9} = \chi_{r1}\chi_{r7}/\chi_{r3}$
N-Y_2'; *C*-Y_2; *P*-Y_1	χ_{r1} (10^3); χ_{r6} (10^2)
N-Y_2'; *C*-Y_2; *P*-Y_1'	χ_{r1} (10^3); χ_{r7} (10^2)
N-Y_2'; *P*-Y_1; *C*-Y_2	χ_{r4} (10^3); $\chi_{r10} = \chi_{r1}\chi_{r6}/\chi_{r4}$
N-Y_2'; *P*-Y_1'; *C*-Y_2	χ_{r5} (10^2); $\chi_{r11} = \chi_{r1}\chi_{r7}/\chi_{r5}$
P-Y_1; *C*-Y_2; *N*-Y_2'	χ_{r2} (10^2); $\chi_{r8} = \chi_{r1}\chi_{r6}/\chi_{r2}$
P-Y_1'; *C*-Y_2; *N*-Y_2'	χ_{r3} (10^3); $\chi_{r9} = \chi_{r1}\chi_{r7}/\chi_{r3}$
P-Y_1; *N*-Y_2'; *C*-Y_2	χ_{r4} (10^3); $\chi_{r10} = \chi_{r1}\chi_{r6}/\chi_{r4}$
P-Y_1'; *N*-Y_2'; *C*-Y_2	χ_{r5} (10^2); $\chi_{r11} = \chi_{r1}\chi_{r7}/\chi_{r5}$

Each row lists one of the 12 ways in which all three receptor-binding domains of Shp2 (*C*, C-SH2; *N*, N-SH2; *P*, PTP) may be sequentially engaged. The receptor site that ultimately engages C-SH2 is denoted Y_2, and Y_1 is the Shp2 substrate site on that receptor; the receptor site that ultimately engages N-SH2 is denoted Y_2', and Y_1' is the Shp2 substrate site on that receptor. The right-hand column specifies the conversion factors χ_r applied to the second and third steps of the complex formation (see also Fig. A4, Appendix A).

considered previously [38], but we arbitrarily assigned one of two different values for each binding mode to illustrate the flexibility of the model. Finally, for simplicity we neglect the formation of chains containing more than one dimer. We found that the large number of combinations of lateral association and ring closure interactions involving such complexes, even when limited to species containing only two dimers, makes the model intractable. Even with all of its limiting assumptions, the model described above is comprised of 149 distinct species participating in 1,032 reactions.

The base model was extended to consider the phosphorylation of Shp2 on a single tyrosine site, which can occur only when Shp2 is receptor-bound. In the phosphorylated state, this site can reversibly engage the N-SH2 domain of the same molecule, provided that the N-

SH2 is free and in the open conformation. This interaction maintains the PTP in the open conformation but also prevents Shp2 association with receptors via N-SH2. Dephosphorylation of the Shp2 phosphorylation site, when not occupied by N-SH2, can occur anywhere in the cell and is modeled simply as a unimolecular transition (Fig. 3.1 *e*). Adding these simple rules increased the complexity of the system by an order of magnitude, yielding a model with 1,325 species and 15,284 reactions.

3.3.2 Implementation of rule-based models

The binding and reaction rules and their associated rate constants were specified in the syntax of the second-generation version of BioNetGen [34], BioNetGen2, which uses graph theoretic methods to automatically generate the associated network of kinetic balances (ordinary differential equations in time). The open-source software (available through http://bionetgen.lanl.gov) uses standard numerical algorithms to solve the generated system of equations, which was deemed to be at steady state at time = 10^3 s. The annotated input files for the base and Shp2 phosphorylation models (Shp2_base.bngl and Shp2_extended.bngl, respectively), which specify the binding/reaction rules are included in Appendix A.

3.3.3 Simplified kinetic model

The simplest model of Shp2/receptor interaction treats SH2 domain-mediated binding of cytosolic Shp2 (*S*) to receptor dimers as a one-step process, with effective forward and reverse rate constants k_+ and k_-, respectively. Shp2 binding is assumed here to be independent of the phosphorylation state of the Shp2 substrate site, which is phosphorylated by the intrinsic kinase with rate constant k_{kin} and, when Shp2 is bound, dephosphorylated with effective rate constant k_{PTP}. *D* and D^* denote receptor dimers free for Shp2 binding and in the unphosphorylated and phosphorylated states, respectively, and *DS* and D^*S denote the corresponding species with Shp2 bound. We define D_{tot} as the total concentration of dimers, according to

$$D_{tot} = D + D^* + DS + D^*S. \tag{3.1}$$

The *bound fraction*, b_D, and *phosphorylated fraction*, p_D are defined as

$$b_D = \frac{DS + D^*S}{D_{tot}}; \quad p_D = \frac{D^* + D^*S}{D_{tot}}. \tag{3.2}$$

At steady state, the simplified model gives the following relationship between p_D and b_D (Eq. A1.8, Appendix A),

$$\begin{aligned} p_D &= \phi\left\{\frac{1 - Qb_D\left[1 + (1-\phi)(1-b_D)\right]}{\phi(1-b_D) + (1-Q)b_D}\right\}; \\ \phi &= \frac{k_{kin}}{k_{kin} + k_{PTP}}; \quad Q = \frac{k_{kin} + k_{PTP}}{k_{\_} + k_{kin} + k_{PTP}}. \end{aligned} \tag{3.3}$$

As in the analysis of related models [40, 41], this relationship is cast in terms of two constant parameters: ϕ, which compares the rates of phosphorylation and dephosphorylation when Shp2 is bound (accordingly, $p_D = \phi$ when $b_D = 1$), and Q, the *exchange quotient*. When $Q \approx 1$ (slow exchange), all dimers not bound to Shp2 remain phosphorylated, whereas a fraction ϕ of the Shp2-bound dimers are phosphorylated, such that $p_D \approx 1 - (1-\phi)b_D$. In the limit of $Q \approx 0$ (rapid exchange), the frequency of dephosphorylation on all receptors approaches $b_D k_{PTP}$, and thus $p_D \approx k_{kin} / (k_{kin} + b_D k_{PTP}) = \phi / [\phi + (1-\phi)b_D]$.

3.4 RESULTS

3.4.1 Shp2 targeting to receptor complexes: multi-valency and serial engagement

The function of Shp2 is quantified in terms of the fraction of receptor sites Y_1 that are phosphorylated (pYR/R_{tot}) at steady state; lower values indicate greater overall rates of Shp2-mediated dephosphorylation. The quasi-steady state approximation is valid when there is a disparity in time scales between that of the intracellular processes (seconds) and that of the binding/trafficking processes that govern the total number of dimerized receptors (minutes), as considered previously [38, 40].

Using the base model, we first systematically varied the concentrations of activated receptors and Shp2 and the binding affinities of the Shp2 SH2 domains (Fig. 3.2). The concentration of receptors in dimers (R_{tot}) was varied in the range of 0.05-0.5 μM (cytosolic volume basis, or $3x10^4$-$3x10^5$ molecules/pL), while the concentration of Shp2 (S_{tot}) was assigned a value of 0.05, 0.1, or 0.2 μM. These receptor and Shp2 concentrations are in the general range of values that are typical of intracellular signaling proteins. Also varied was the dissociation rate constant k_{off} of the SH2 domain-receptor interactions, assumed to be equivalent for the C-SH2 and N-SH2 domains. The values are such that the single-site dissociation constants, K_D, are in the range of 0.1-10 μM, which spans the range of highly specific and regulatory SH2-phosphotyrosine interactions [20, 42]. As expected, essentially all dimers are bound with Shp2 when the affinity is sufficiently high (low k_{off}) and Shp2 outnumbers activated receptors ($R_{tot}/2S_{tot} < 1$); this yields the minimum extent of receptor phosphorylation (Fig. 3.2). As the SH2 domain affinities are decreased, Shp2-receptor binding decreases dramatically, because both SH2 domains are affected.

The fraction of phosphorylated receptors tends to increase as the number of activated receptors is increased, which saturates the binding of Shp2. When activated receptors greatly outnumber Shp2 molecules, a higher dissociation rate from receptors can actually be advantageous. This is a manifestation of the *serial engagement* effect, more commonly associated with T-cell receptor activation [43-45], whereby one molecule of Shp2 encounters and dephosphorylates multiple dimers before the kinase activity can restore their phosphorylation. Indeed, model calculations show that, when activated receptors are

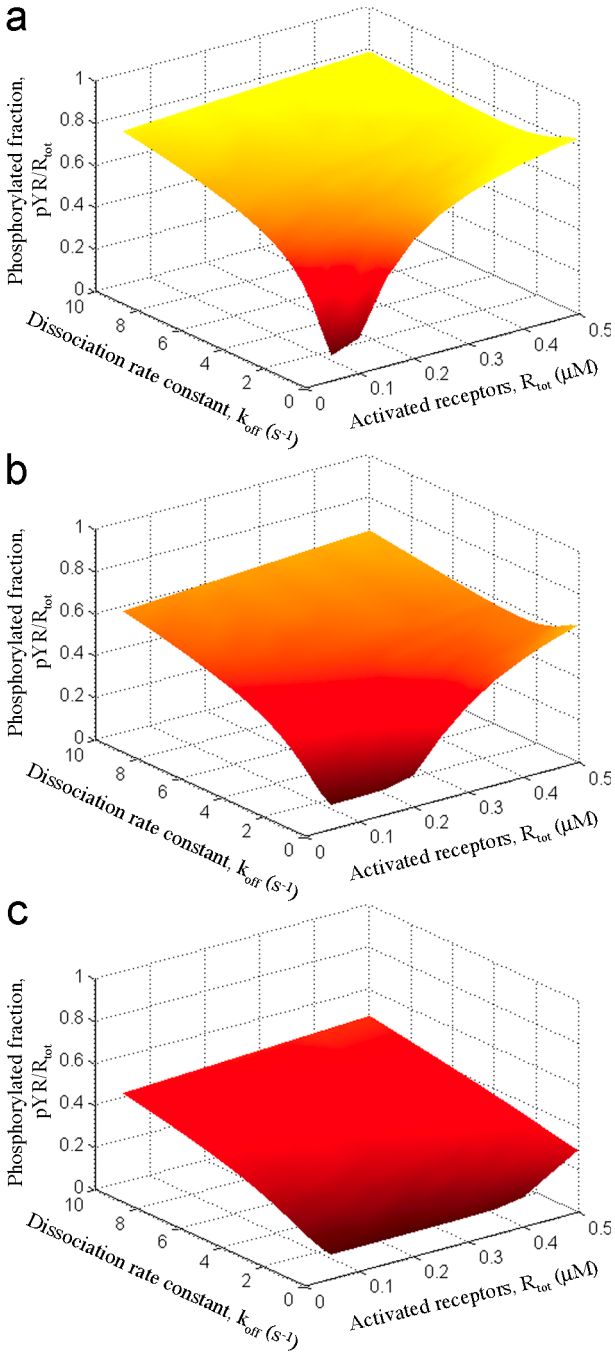


FIGURE 3.2 Steady-state analysis of Shp2-mediated receptor dephosphorylation. The fraction of activated receptors phosphorylated on the Shp2 substrate site, Y_1 (pYR/R_{tot}), was calculated as a function of the abundance of activated receptors, mimicking a dose response or varying expression levels (R_{tot}), and the binding affinities of the two Shp2 SH2 domains (varied in terms of the dissociation rate constant, $k_{off,CSH2} = k_{off,NSH2} = k_{off}$). The other variable is the abundance of Shp2 (S_{tot}): (*a*) S_{tot} = 0.05 μM; (*b*) S_{tot} = 0.10 μM; (*c*) S_{tot} = 0.20 μM. All other rate constants are as listed in Tables 3.1 and 3.2.

relatively abundant, there is a value of k_{off} that maximizes the extent of receptor dephosphorylation (Fig. 3.2 *a* & *b*; high R_{tot}, $k_{off} \approx$ 1-3 s^{-1}).

The serial engagement effect can be understood in the context of a simple kinetic model, in which the association and dissociation of Shp2 from receptor complexes and Shp2-mediated dephosphorylation of receptors are approximated as single steps (Fig. 3.3). In this model, the potential for serial engagement is encapsulated in a single dimensionless parameter, the exchange quotient Q that characterizes the number of phosphorylation and dephosphorylation reactions that occur during the lifetime of an Shp2/receptor encounter (Eq. 3.3, Methods). When $Q \approx 1$, all activated receptors not bound to Shp2 are maintained in a fully phosphorylated state, whereas a value of Q significantly less than 1 is indicative of serial engagement and its degree. As a function of the fraction of dimers bound, b_D, a lower value of Q always improves the overall rate of dephosphorylation. This is evident when the results from Fig. 3.2 are plotted in this way and compared with the predictions of the simplified kinetic model; increasing the SH2 domain dissociation constant decreases the average lifetime of the Shp2/receptor association, decreasing the apparent value of Q and enhancing serial engagement (Fig. 3.3).

Intra-complex binding, controlled by the magnitudes of the conversion factors χ_r, is critical for high-avidity binding of Shp2 (Fig. 3.4). Increasing the rates of intra-complex binding affects Shp2 function in the same manner as decreasing k_{off} and can offset a reduction

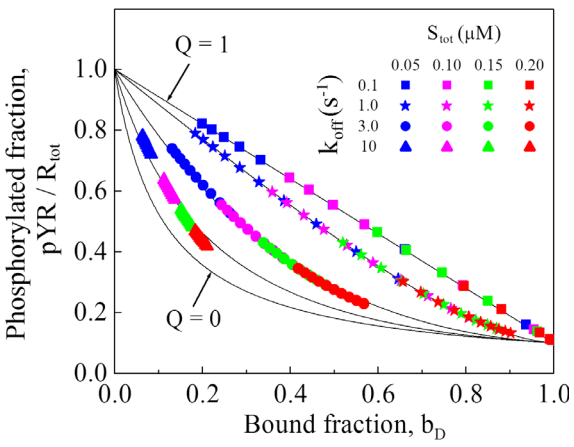


FIGURE 3.3 Serial engagement of activated receptors by Shp2. The receptor Y_1 phosphorylation results shown in Fig. 3.2 are plotted as a function of the fraction of receptor dimers bound with at least one Shp2 (b_D). Plotted in this way, the results may be compared with the predictions of a simplified, two-parameter model (Methods, Eq. 3.3 and Appendix A, Eq. A1.8. One parameter in this model, ϕ, was determined from the value of pYR/R_{tot} at $b_D \approx 1$. The value of the other, the exchange quotient Q, determines the degree of serial engagement, the ability of an Shp2 molecule to engage and dephosphorylate numerous receptor complexes for each instance of Y_1 phosphorylation (solid curves). The values of Q, in descending order, are 1, 0.97, 0.81, 0.51, and 0.

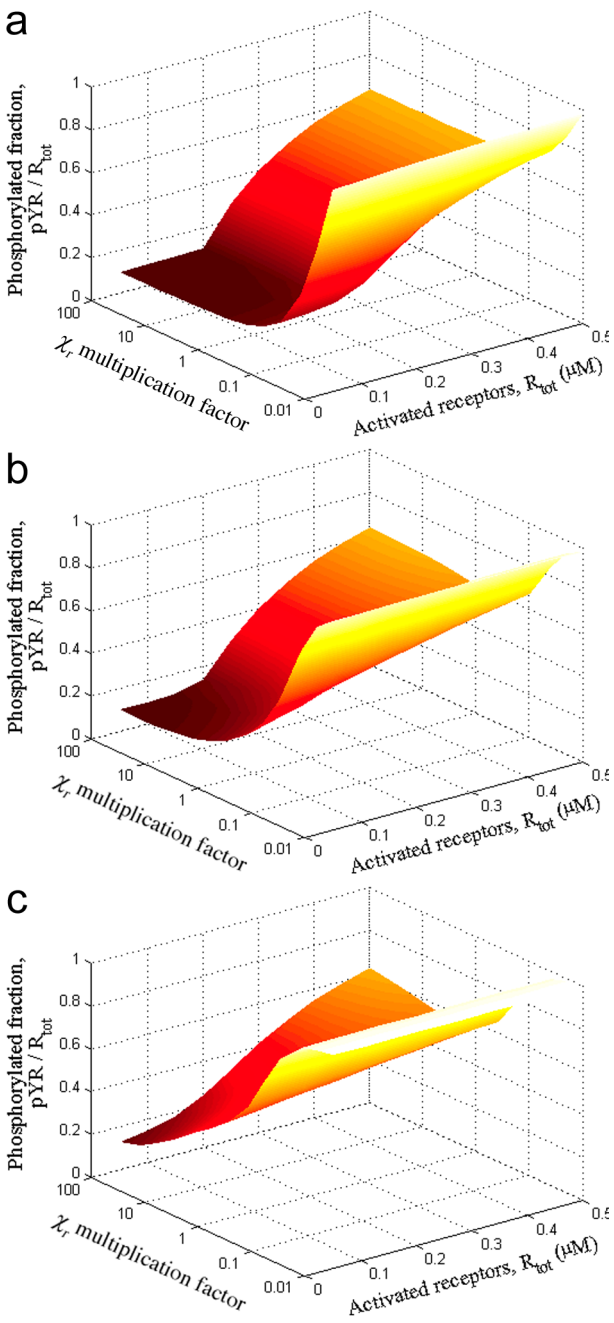


FIGURE 3.4 Intra-complex binding is important for high-avidity binding of Shp2. The fraction of activated receptors phosphorylated on the Shp2 substrate site, Y_1 (pYR/R_{tot}), was calculated as a function of the abundance of activated receptors (R_{tot}) and the intra-complex association rate constants (varied by multiplying the base-case values of χ_r, listed in Table 3.2, by a common factor). The value of $k_{off,CSH2} = k_{off,NSH2} = k_{off}$ was varied according to (*a*) $k_{off} = 0.1$ s^{-1}; (*b*) $k_{off} = 1.0$ s^{-1}; (*c*) $k_{off} = 5.0$ s^{-1}. In all cases, $S_{tot} = 0.10$ μM, and other parameter values are as listed in Table 3.1.

in the affinities of the individual SH2 domains. Both parameters influence the average rate of Shp2 dissociation from the receptor complex, which was taken as a lumped parameter in our simplified kinetic model (Eq. 3.3). Thus, when k_{off} is low or χ_r are high, Shp2 binding to

receptors is approximately stoichiometric (Fig. 3.4 *a*). When k_{off} is higher and χ_r are modest, Shp2 binds with lower avidity, and there is also the opportunity for serial engagement when activated receptors outnumber Shp2 molecules (Fig. 3.4 *c*).

3.4.2 Deletion of the Shp2 N-SH2 domain can diminish or enhance Shp2 function; C-terminal phosphorylation of Shp2 mimics N-SH2 deletion

Based on dephosphorylation of generic phosphorylated substrates in solution, it is clear that deletion of the N-SH2 domain releases the auto-inhibition of the PTP catalytic site, yielding full catalytic activity of the Shp2 enzyme. Based on the cooperativity of the two SH2 domains in targeting Shp2 to activated receptors and other signaling complexes at the plasma membrane, however, it is not immediately apparent how this modification would affect dephosphorylation of substrates associated with those complexes and thus modulation of signal transduction by Shp2 in cells.

Indeed, model calculations show that deletion of the N-SH2 can either diminish or enhance the dephosphorylation of a targeted substrate such as a receptor phosphorylation site (Fig. 3.5). The absence of the N-SH2 domain enhances the activity of Shp2 while associated with receptors through its C-SH2 domain, but at the same time it affects the avidity of the Shp2/receptor interaction. The calculations show that, under the conditions where activated receptors are predominantly associated with full-length Shp2, N-SH2 deletion tends to diminish substrate dephosphorylation; conversely, when either the activated receptors are in excess over Shp2, or the SH2 domain/receptor interactions yield a relatively low avidity, N-SH2 deletion can enhance substrate dephosphorylation (Fig. 3.5 *a*-*c*). Based on the concepts established in the previous section, it is clear that there are two distinct effects that contribute to the enhancement offered by N-SH2 deletion: the lack of PTP auto-inhibition, and the serial engagement of receptors as the lifetime of Shp2/receptor complexes is decreased.

To further illustrate the importance of the parameter values on the predicted effect of N-SH2 deletion in Shp2 signaling, the SH2 domain affinities were adjusted to a relatively low value (K_D = 5 μM, as in Fig. 3.5 *c*), while both increasing the catalytic rates of receptor Y_1 phosphorylation and dephosphorylation by ten-fold (to negate the serial engagement

effect) and increasing the rates of all intra-complex associations by 100-fold (Fig. 3.5 *d*); under these conditions, the full-length Shp2 is capable of high-avidity binding to activated receptors, even though the affinities of the individual SH2 domains are low. It is clear that N-

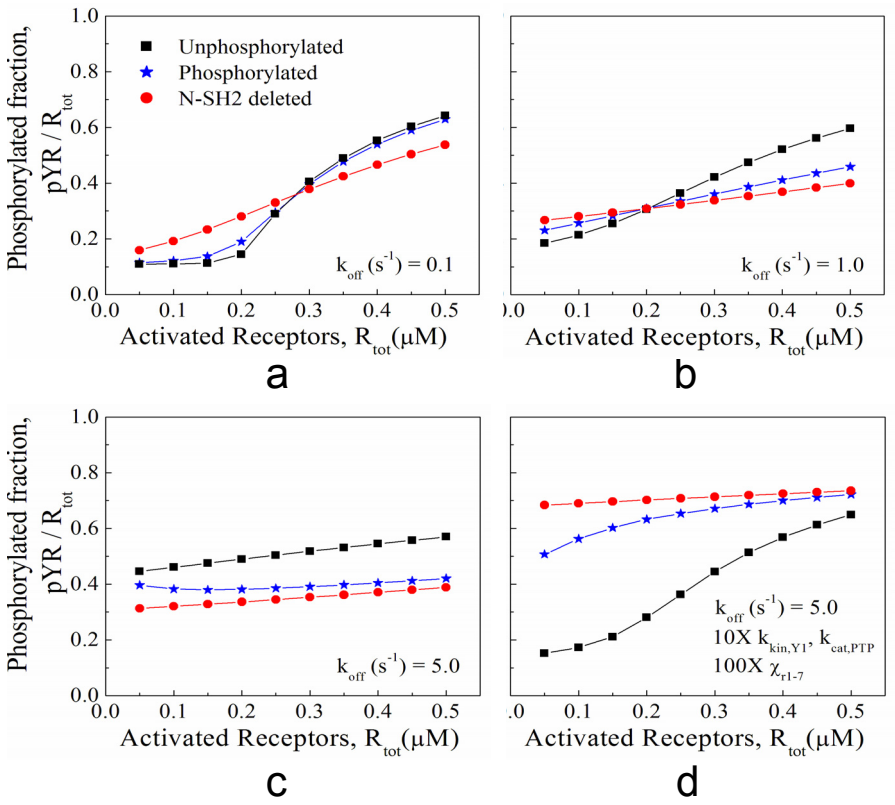


FIGURE 3.5 Conflicting roles of the N-SH2 domain in Shp2 auto-inhibition and receptor targeting. The base model was modified to exclude all binding interactions of the N-SH2 domain, and the function of N-SH2-deleted Shp2 (red circles) was compared with full-length Shp2 (black squares) for various levels of activated receptors and an Shp2 expression level of S_{tot} = 0.10 μM. An extended model shows that receptor-mediated phosphorylation of Shp2, which occupies the N-SH2 in an intramolecular fashion, yields an intermediate level of receptor dephosphorylation (blue stars). The SH2 dissociation rate constant, $k_{off,CSH2}$ = $k_{off,NSH2}$ = k_{off} was varied according to: (*a*) k_{off} = 0.1 s^{-1}; (*b*) k_{off} = 1.0 s^{-1}; (*c* and *d*), k_{off} = 5.0 s^{-1}. Other rate constant values are as listed in Tables 3.1 and 3.2. In D, the values of $k_{kin,Y1}$ and $k_{cat,PTP}$ were increased 10-fold, and all intra-complex association rates were increased 100-fold.

SH2 deletion compromises Shp2 function under these conditions, particularly when Shp2 molecules outnumber activated receptors.

As illustrated in Fig. 3.1 *e*, the base model was extended to include phosphorylation of receptor-bound Shp2 on its C-terminal tail; this phosphorylation site may then engage the

N-SH2 domain, which prevents N-SH2 from either auto-inhibiting the PTP domain or participating in receptor binding. While engaged in this manner, Shp2 functions as it does when N-SH2 is deleted altogether. Accordingly, the output of the extended model with Shp2 phosphorylation is invariably bracketed by the results of the base model for full-length and N-SH2-deleted Shp2 (Fig. 3.5), depending on the relative extent of Shp2 phosphorylation and whether the Shp2 phosphorylation site or activated receptors better compete for N-SH2 binding. The model therefore predicts that, in the same fashion as N-SH2 deletion, tyrosine phosphorylation of Shp2 can either diminish or enhance Shp2-mediated dephosphorylation of receptor-associated substrates.

3.4.2 Deletion of the Shp2 C-SH2 domain, or of both SH2 domains, compromises Shp2 function

With its C-SH2 domain deleted, Shp2 remains auto-inhibited by its N-SH2, and any receptor engagement must occur while Shp2 is in the open conformation. Even with the SH2 domain K_D skewed towards the highest affinity (0.1 μM, comparable with the concentrations of Shp2 and activated receptors), the auto-inhibition drastically reduces the rate of Shp2-receptor association when C-SH2 is absent. Thus, only a small fraction of activated receptors are associated with and dephosphorylated by C-SH2-deleted Shp2 (Fig. 3.6 *a*). For lower SH2 domain affinities (K_D up to 10 μM) as well, it was confirmed that C-SH2 deletion always diminishes receptor dephosphorylation, despite the greater degree of serial engagement. If one swaps the frequencies of the open and close transitions, such that the open conformation is highly favored, the performance is nearly identical to the N-SH2 deletion mutant (Fig. 3.6 *a* and results not shown).

Like the C-SH2 deletion, truncation of both SH2 domains yields an Shp2 variant that is deficient in dephosphorylating receptors, but for a different reason. Whereas the C-SH2 deletion has reduced targeting capability and full auto-inhibition of the PTP, deletion of both SH2 domains yields a fully active enzyme in solution but erases all targeting capability in cells. Which of these deficiencies is more debilitating for Shp2 function depends on the binding and catalytic properties of the enzyme (Fig. 3.6 *b*). For the set of parameter values

considered here, the catalytic efficiency of the SH2-null variant is sensitive to a change in the catalytic rate constant, $k_{cat,PTP}$. In contrast, for C-SH2-deleted Shp2, the cooperative binding of the N-SH2 and PTP domains gives rise to a scenario in which the lifetime of the Shp2/receptor complex is essentially limited by the PTP catalytic step, and the extent of dephosphorylation is thus determined by the rate of Shp2-receptor association.

3.4.3 Evaluation of restrictive binding rules for PTP-substrate binding within the complex

Our model considers all of the possible ways in which the two SH2 domains and PTP domain of Shp2 may participate in interactions with activated receptors. The multiplication factors for intra-complex binding, χ_r, were arbitrarily assigned different values to illustrate the generality of the model (Table 3.2), but currently there is no reason to suspect that certain modes of intra-complex binding would be favored or disfavored. One can, however, assess the possible impact of structural constraints on the dephosphorylation of receptor-associated substrates (Fig. 3.7).

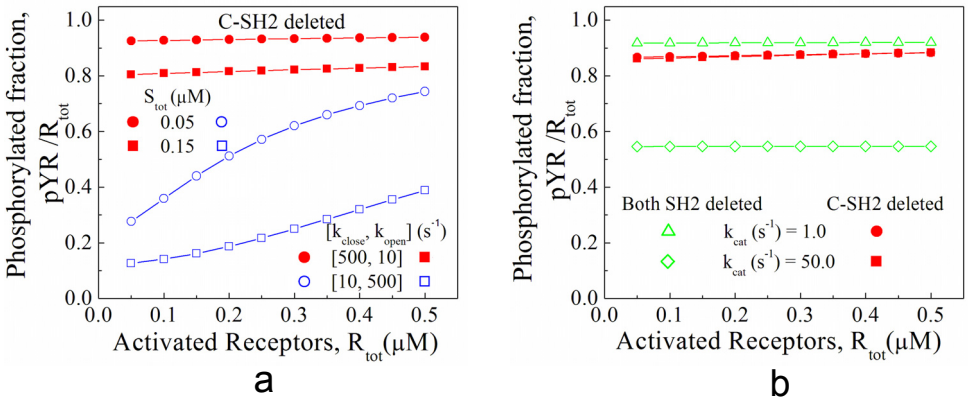


FIGURE 3.6 The C-SH2 domain is essential for receptor targeting of Shp2. The base model was modified to exclude all binding interactions of the C-SH2 domain, or of both the C-SH2 and N-SH2 domains. Receptor dephosphorylation was assessed for various levels of activated receptors and $k_{off,NSH2}$ = 0.1 s^{-1}. Except where indicated, all other rate constant values are as listed in Tables 3.1 & 3.2. (*a*) Defective Shp2 function with C-SH2 deleted (closed symbols). Shp2 concentration S_{tot} values are as indicated. Shift of the open/close equilibrium to favor the open conformation (open symbols) mimics the N-SH2 deletion case. (*b*) Comparison of C-SH2 deletion (closed symbols) with deletion of both SH2 domains (open symbols), for different values of the PTP catalytic rate constant, k_{cat}.

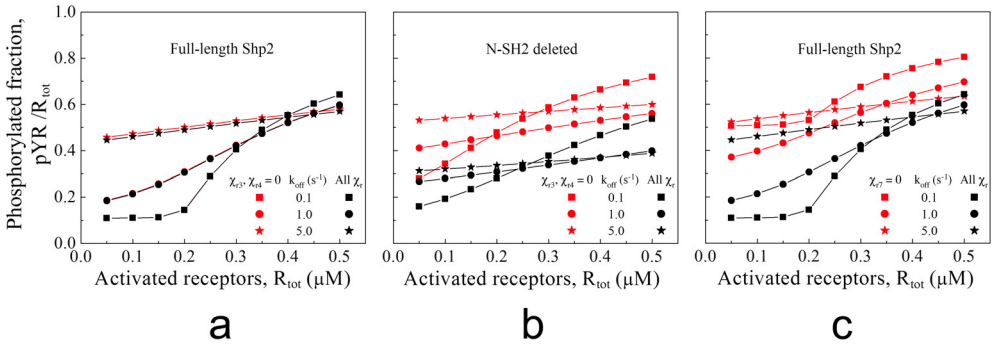


FIGURE 3.7 Potential restrictions on PTP-substrate binding within the complex. Certain of the multiplication factors χ_r were assigned either base values (Table 3.2; black) or a value of zero (red) to assess the corresponding modes of PTP-Y_1 interaction after binding of one or both SH2 domains of Shp2: (*a*) $\chi_{r3} = \chi_{r4} = 10^3$ or 0 μM; (*b*) same as *a* but with N-SH2-deleted Shp2; (*c*) $\chi_{r7} = 100$ or 0 μM. In all cases, $S_{tot} = 0.10$ μM, and SH2 dissociation rate constant ($k_{off,CSH2} = k_{off,NSH2} = k_{off}$) values are as indicated.

For instance, if Shp2 with only one SH2 domain bound is to dephosphorylate the receptor, the PTP may be restricted to binding only one of the two receptors in the dimer. This is of little consequence for full-length Shp2, because simultaneous binding of both SH2 domains is greatly favored (Fig. 3.7 *a*). On the other hand, when one of the SH2 domains is deleted, this restriction diminishes receptor dephosphorylation significantly; under these conditions, N-SH2 deletion would be less likely to enhance Shp2 function relative to the full-length enzyme (Fig. 3.7 *b*). Conversely, when both of its SH2 domains are bound, full-length Shp2 is compromised if its PTP domain is constrained to binding only one of the receptors (e.g., whichever receptor is bound to C-SH2, Fig. 3.7 *c*), whereas this restriction is not applicable to N-SH2-deleted Shp2. Given the dynamic interactions between full-length Shp2 and a receptor dimer, it is still possible in this case for the PTP domain to interact with the substrate sites of both receptors before dissociating with the complex completely, but it does so with greatly reduced proclivity.

3.5 DISCUSSION

The function of a signaling protein is generally pieced together from structural analysis, through the identification of homologous domains and their arrangement in the protein's tertiary structure, and biochemical characterization of binding specificities and affinities in solution. In the case of Shp2, an SH2 domain-containing PTP, a consistent story has thus emerged. Full-length Shp2 tends to adopt a closed, auto-inhibited conformation that blocks the active site of the PTP; stabilization of the fully active, open conformation of Shp2 requires the occupancy of its N-SH2 domain, either by tyrosine-phosphorylated proteins/peptides or by one of two tyrosine phosphorylation sites in the C-terminal tail of Shp2 itself. The function of Shp2 in cells adds another dimension of complexity, namely that the SH2 domains are also important for targeting Shp2 to signaling complexes, where the PTP enjoys an induced proximity effect. To synthesize and analyze such dynamics in signal transduction, kinetic modeling has become an increasingly valuable tool, but the numerous combinations of binding and phosphorylation states imposes a significant technical challenge. The ruled-based approach employed here permitted us to examine how multiple protein domains with different functionalities work in concert to affect signaling function.

The central feature of our model is the cooperative binding of the two SH2 domains of Shp2. We assume binding to symmetric phosphorylation sites on a pair of dimerized receptors, but the conclusions are expected to be sufficiently general. Experiments with individual or tandem SH2 domains of Shp2 and singly or doubly phosphorylated peptides show that the interaction of both SH2 domains, which individually possess low single-site affinities (K_D ~ 1-10 μM), yields a high overall binding avidity (effective K_D ~ 1-10 nM) that could be sufficient for near stoichiometric binding in cells [18-20]. A 1,000-fold enhancement in binding translates to a value of χ_r that is roughly 1,000 times higher than the single-site K_D, or χ_r ~ 1-10 mM, in line with our conservative estimates. In the context of full-length Shp2, to the extent that N-SH2 cannot bind to phosphotyrosine while in the closed conformation, the value of χ_r would need to be even higher to overcome this effect (see Figs. 3.4 *c* and 3.5 *d*). When Shp2 binding avidity is not as strong, we found that the effect of

serial engagement can at least partially compensate; this effect arises when encounters between Shp2 and receptors are fleeting relative to the rate of substrate phosphorylation/dephosphorylation (Fig. 3.3).

The disparity between affinity and avidity of the Shp2 SH2 domains is at the core of what is perhaps our most interesting finding, that deletion of the N-SH2 domain can either diminish or enhance receptor dephosphorylation (Fig. 3.5). Under conditions that favor cooperative binding of a high fraction of receptors, N-SH2 deletion tends to abrogate Shp2 function; otherwise, Shp2 function may be enhanced. Interestingly, the effect of N-SH2 deletion can be reversed as the number of activated receptors is modulated, as by a change in growth factor/cytokine concentration or by receptor overexpression/knock-down.

These predictions of the model can be related to published data, although it should be noted that those data do not allow for a quantitative comparison. Targeted deletion of Shp2 exon 3 in mouse yields expression of a mutant Shp2 with most of its N-SH2 domain deleted. As expected, this mutation yields greater overall PTP activity as measured in solution, but the Shp2 variant is severely defective in binding to activated receptors and other signaling complexes [46-50]. The effect of this mutation on intracellular signal transduction was assessed at the level of growth factor-stimulated Erk activation, which is positively modulated by Shp2, with variable results. In one study, Erk phosphorylation stimulated by PDGF was enhanced in the mutant cells relative to wild-type, despite a lower expression level (~ 25% of wild-type) of the mutant, whereas Erk signaling stimulated by fibroblast growth factor was significantly reduced in the same cells [46]. In another study, Erk signaling stimulated by all growth factors tested, including PDGF, was diminished in mutant-expressing cells [48]. These observations are consistent with our modeling results, which predict that different receptor expression levels and/or Shp2 binding avidities can affect the outcome of N-SH2 deletion. Alternatively, some receptors may activate Shp2 for dephosphorylating substrates not intimately associated with the complex, in which case targeting of Shp2 with N-SH2 deleted is not necessary.

By comparison with N-SH2 deletion, certain Shp2 mutations disrupt auto-inhibition of the PTP without preventing binding of phosphorylated peptides to N-SH2; indeed,

favoring the open conformation enhances this function. Interestingly, such activating Shp2 mutation sites are associated with Noonan's syndrome, a common human birth defect [51]. Expression of Shp2 variants activated in this manner generally yields enhanced growth factor stimulation of Erk signaling, at least in certain cell contexts, arguing for the importance of the N-SH2 targeting function [52-56].

Our modeling approach also sheds light on the most controversial aspect of Shp2 regulation, the mechanism by which Shp2 phosphorylation modulates its activity and signaling functions. In our extended model, Shp2 phosphorylation at one site and the intramolecular binding of N-SH2 were considered. As one might have predicted, this gives a level of receptor dephosphorylation that lies between those mediated by full-length, unphosphorylated Shp2 and Shp2 with N-SH2 deleted (Fig. 3.5). In other words, intramolecular binding of N-SH2 mimics N-SH2 deletion in our model. What if we had included both of the Shp2 phosphorylation sites? Biochemical evidence suggests that these sites can engage both SH2 domains, favoring the open conformation of Shp2 in a cooperative manner [26]. If so, it is clear that inclusion of both phosphorylation sites would mimic the deletion of both SH2 domains; such a mechanism discounts any targeting function of the SH2 domains. Alternative mechanisms for the effect(s) of Shp2 phosphorylation include binding of the adaptor protein Grb2, which might help target Shp2 to Gab1 in signaling by EGF receptor and other RTKs [57], and binding of SH2 domain-containing substrates of Shp2 [24].

This modeling study underscores the complexity of interactions between signaling proteins with multiple modular domains. In the case of Shp2, it illustrates the potential trade-offs between regulation of catalytic activity and targeting of the enzyme to substrate-containing complexes or compartments, and it shows that intracellular Shp2 and receptor expression levels must be carefully considered in the interpretation of cell signaling experiments. Conversely, these expression levels would need to be measured and varied systematically if the quantitative predictions of the model are to be validated. Rule-based kinetic modeling is a powerful computational tool for modeling the assembly of signaling

complexes and signal transduction pathways, and here we have shown how it can be used to impart structure-based functionality to their molecular components.

3.6 REFERENCES

1. NK Tonks, BG Neel: Combinatorial control of the specificity of protein tyrosine phosphatases. *Curr. Opin. Cell Biol.* 2001, 13:182-195.

2. P van der Geer, T Hunter, RA Lindberg: Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu. Rev. Cell Biol.* 1994, 10:251-337.

3. J Schlessinger: Cell signaling by receptor tyrosine kinases. *Cell* 2000, 103:211-225.

4. T Hunter: Signaling– 2000 and beyond. *Cell* 2000, 100:113-127.

5. T Pawson: Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* 2004, 116:191-203.

6. RP Bhattacharyya, A Remenyi, BJ Yeh, WA Lim: Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* 2006, 75:655-680.

7. BG Neel, HH Gu, L Pao: The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. *Trends Biochem. Sci.* 2003, 28:284-293.

8. RA Klinghoffer, A Kazlauskas: Identification of a putative Syp substrate, the PDGFβ receptor. *J. Biol. Chem.* 1995, 270:22208-22217.

9. S Ekman, A Kallin, A Engström, C-H Heldin: SHP-2 is involved in heterodimer specific loss of phosphorylation of Tyr771 in the PDGF β-receptor. *Oncogene* 2002, 21:1870-1875.

10. YM Agazie, MJ Hayman: Molecular mechanism for a role of SHP2 in epidermal growth factor receptor signaling. *Mol. Cell. Biol.* 2003, 23:7875-7886.

11. P Hof, S Pluskey, S Dhe-Paganon, MJ Eck, SE Shoelson: Crystal structure of the tyrosine phosphatase SHP-2. *Cell* 1998, 92:441-450.

12. D Barford, BG Neel: Revealing mechanisms for SH2 domain mediated regulation of the protein tyrosine phosphatase SHP-2. *Structure* 1998, 6:249-254.

13. RJ Lechleider, S Sugimoto, AM Bennett, AS Kashishian, JA Cooper, SE Shoelson, CT Walsh, BG Neel: Activation of the SH2-containing phosphotyrosine phosphatase SH-PTP2 by its binding site, phosphotyrosine 1009, on the human platelet-derived growth factor receptor. *J. Biol. Chem.* 1993, 268:21478-21481.

14. U Dechert, M Adam, KW Harder, I Clark-Lewis, F Jirik: Characterization of protein tyrosine phosphatase SH-PTP2: study of phosphopeptide substrates and possible regulatory role of SH2 domains. *J. Biol. Chem.* 1994, 269:5602-5611.

15. S Sugimoto, TJ Wandless, SE Shoelson, BG Neel, CT Walsh: Activation of the SH2-containing protein tyrosine phosphatase, SH-PTPs by phosphotyrosine-containing peptides derived from insulin receptor substrate-1. *J. Biol. Chem.* 1994, 269:13614-13622.

16. DH Pei, J Wang, CT Walsh: Differential functions of the two Src homology 2 domains in protein tyrosine phosphatase SH-PTP1. *Proc. Natl. Acad. Sci. U.S.A.* 1996, 93:1141-1145.

17. H Keilhack, FS David, M McGregor, LC Cantley, BG Neel: Diverse biochemical properties of Shp2 mutants: implications for disease phenotypes. *J. Biol. Chem.* 2005, 280:30984-30993.

18. S Pluskey, TJ Wandless, CT Walsh, SE Shoelson: Potent stimulation of SHPTPs phosphatase activity by simultaneous occupancy of both SH2 domains. *J. Biol. Chem.* 1995, 270:2897-2900.

19. MJ Eck, S Pluskey, T Trub, SC Harrison, SE Shoelson: Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. *Nature* 1996, 379:277-280.

20. EA Ottinger, MC Botfield, SE Shoelson: Tandem SH2 domains confer high specificity in tyrosine kinase signaling. *J. Biol. Chem.* 1998, 273:729-735.

21. JM Cunnick, L Mei, CA Doupnik, J Wu: Phosphotyrosines 627 and 659 of Gab1 constitute a bisphosphoryl tyrosine-based activation motif (BTAM) conferring binding and activation of SHP2. *J. Biol. Chem.* 2001, 276:24380-24387.

22. G Feng, C Hui, T Pawson: SH2-containing phosphotyrosine phosphatase as a target of protein-tyrosine kinases. *Science* 1993, 259:1607-1610.

23. W Vogel, R Lammers, J Huang, A Ullrich: Activation of a phosphotyrosine phosphatase by tyrosine phosphorylation. *Science* 1993, 259:1611-1614.

24. T Araki, H Nawa, BG Neel: Tyrosyl phosphorylation of Shp2 is required for normal ERK activation in response to some, but not all, growth factors. *J. Biol. Chem.* 2003, 278:41677-41684.

25. W Lu, DQ Gong, D Bar-Sagi, PA Cole: Site-specific incorporation of a phosphotyrosine mimetic reveals a role for tyrosine phosphorylation of SHP-2 in cell signaling. *Mol. Cell* 2001, 8:759-769.

26. ZS Zhang, K Shen, W Lu, PA Cole: The role of C-terminal tyrosine phosphorylation in the regulation of SHP-1 explored via expressed protein ligation. *J. Biol. Chem.* 2003, 278:4668-4674.

27. W Lu, K Shen, PA Cole: Chemical dissection of the effects of tyrosine phosphorylation of SHP-2. *Biochemistry* 2003, 42:5461-5468.

28. AM Bennett, TL Tang, S Sugimoto, CT Walsh, BG Neel: Protein-tyrosine-phosphatase SHPTP2 couples platelet-derived growth factor receptor β to Ras. *Proc. Natl. Acad. Sci. U.S.A.* 1994, 91:7335-7339.

29. W Li, R Nishimura, A Kashishian, AG Batzer, WJH Kim, JA Cooper, J Schlessinger: A new function for a phosphotyrosine phosphatase: linking Grb2-Sos to a receptor tyrosine kinase. *Mol. Cell. Biol.* 1994, 14:509-517.

30. G Weng, US Bhalla, R Iyengar: Complexity in biological signaling systems. *Science* 1999, 284:92-96.

31. BN Kholodenko: Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell. Biol.* 2006, 7:165-176.

32. WS Hlavacek, JR Faeder, ML Blinov, AS Perelson, B Goldstein: The complexity of complexes in signal transduction. *Biotechnol. Bioeng.* 2003, 84:783-794.

33. JR Faeder, WS Hlavacek, I Reischl, ML Blinov, H Metzger, A Redondo, C Wofsy, B Goldstein: Investigation of early events in FcεRI-mediated signaling using a detailed mathematical model. *J. Immunol.* 2003, 170:3769-3781.

34. ML Blinov, JR Faeder, B Goldstein, WS Hlavacek: BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* 2004, 20:3289-3291.

35. L Lok, R Brent: Automatic generation of cellular reaction networks with Moleculizer 1.0. *Nat. Biotechnol.* 2005, 23:131-136.

36. H Conzelmann, J Saez-Rodriguez, T Sauter, BN Kholodenko, ED Gilles: A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks. *BMC Bioinformatics* 2006, 7:34.

37. WS Hlavacek, JR Faeder, ML Blinov, RG Posner, M Hucka, W Fontana: Rules for modeling signal-transduction systems. *Sci. STKE* 2006, 344:re6.

38. JM Haugh, IC Schneider, JM Lewis: On the cross-regulation of protein tyrosine phosphatases and receptor tyrosine kinases in intracellular signaling. *J. Theor. Biol.* 2004, 230:119-132.

39. A Kazlauskas, GS Feng, T Pawson, M Valius: The 64-kDa protein that associates with the platelet-derived growth factor receptor β subunit via Tyr-1009 is the SH2-containing phosphotyrosine phosphatase Syp. *Proc. Natl. Acad. Sci. U.S.A.* 1993, 90:6939-6942.

40. JM Haugh, DA Lauffenburger: Analysis of receptor internalization as a mechanism for modulating signal transduction. *J. Theor. Biol.* 1998, 195:187-218.

41. JM Haugh, AC Huang, HS Wiley, A Wells, DA Lauffenburger: Internalized epidermal growth factor receptors participate in the activation of $p21^{ras}$ in fibroblasts. *J. Biol. Chem.* 1999, 274:34350-34360.

42. JE Ladbury, MA Lemmon, M Zhou, J Green, MC Botfield, J Schlessinger: Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. *Proc. Natl. Acad. Sci. U.S.A.* 1995, 92:3199-3203.

43. S Valitutti, S Muller, M Cella, E Padovan, A Lanzavecchia: Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* 1995, 375:148-151.

44. S Valitutti, A Lanzavecchia: Serial triggering of TCRs: a basis for the sensitivity and specificity of antigen recognition. *Immunol. Today* 1997, 18:299-304.

45. D Coombs, AM Kalergis, SG Nathenson, C Wofsy, B Goldstein: Activated TCRs remain marked for internalization after dissociation from pMHC. *Nat. Immunol.* 2002, 3:926-931.

46. TM Saxton, M Henkemeyer, S Gasca, R Shen, DJ Rossi, F Shalaby, G Feng, T Pawson: Abnormal mesoderm patterning in mouse embryos mutant for the SH2 tyrosine phosphatase Shp-2. *EMBO J.* 1997, 16:2352-2364.

47. C Qu, Z Shi, R Shen, F Tsai, SH Orkin, G Feng: A deletion mutation in the SH2-N domain of Shp-2 severely suppresses hematopoietic cell development. *Mol. Cell. Biol.* 1997, 17:5499-5507.

48. Z Shi, W Lu, G Feng: The Shp-2 tyrosine phosphatase has opposite effects in mediating the activation of extracellular signal-regulated and c-Jun NH_2-terminal mitogen-activated protein kinases. *J. Biol. Chem.* 1998, 273:4904-4908.

49. Z Shi, D Yu, M Park, M Marshall, G Feng: Molecular mechanism for the Shp-2 tyrosine phosphatase function in promoting growth factor stimulation of Erk activity. *Mol. Cell. Biol.* 2000, 20:1526-1536.

50. W Yang, LD Klaman, B Chen, T Araki, H Harada, SM Thomas, EL George, BG Neel: An Shp2/SFK/Ras/Erk signaling pathway controls trophoblast stem cell survival. *Dev. Cell* 2006, 10:317-327.

51. M Tartaglia, EL Mehler, R Goldberg, G Zampino, HG Brunner, H Kremer, I van der Burgt, AH Crosby, A Ion, S Jeffery, et al: Mutations in *PTPN11*, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.* 2001, 29:465-468.

52. AM O'Reilly, S Pluskey, SE Shoelson, BG Neel: Activated mutants of SHP-2 preferentially induce elongation of *Xenopus* animal caps. *Mol. Cell. Biol.* 2000, 20:299-311.

53. T Araki, MG Mohi, FA Ismat, RT Bronson, IR Williams, JL Kutok, WT Yang, LI Pao, DG Gilliland, JA Epstein, et al: Mouse model of Noonan syndrome reveals cell type- and gene dosage-dependent effects of Ptpn11 mutations. *Nat. Med.* 2004, 10:849-857.

54. MG Mohi, IR Williams, CR Dearolf, G Chan, JL Kutok, S Cohen, K Morgan, C Boulton, H Shigematsu, H Keilhack, et al: Prognostic, therapeutic, and mechanistic implications of a mouse model of leukemia evoked by Shp2 (PTPN11) mutations. *Cancer Cell* 2005, 7:179-191.

55. RJ Chan, MB Leedy, V Munugalavadia, CS Voorhorst, YJ Li, MG Yu, R Kapur: Human somatic *PTPN11* mutations induce hematopoietic-cell hypersensitivity to granulocyte-macrophage colony-stimulating factor. *Blood* 2005, 105:3737-3742.

56. WM Yu, H Daino, J Chen, KD Bunting, CK Qu: Effects of a leukemia-associated gain-of-function mutation of SHP-2 phosphatase on interleukin-3 signaling. *J. Biol. Chem.* 2006, 281:5426-5434.

57. HH Gu, BG Neel: The 'Gab' in signal transduction. *Trends Cell Biol.* 2003, 13:122-130.

58. SH Northrup, HP Erickson: Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U.S.A.* 1992, 89:3338-3342.

59. S Sugimoto, RJ Lechleider, SE Shoelson, BG Neel, CT Walsh: Expression, purification, and characterization of SH2-containing protein tyrosine phosphatase, SH-PTP2. *J. Biol. Chem.* 1993, 268:22771-22776.

CHAPTER 4

Computational models of tandem Src homology 2 domain interactions and application to phosphoinositide 3-kinase

Adapted from Barua, Faeder and Haugh, J. Biol. Chem., 283: 7338 – 7345 (2008)

4.1 ABSTRACT

Intracellular signal transduction proteins typically utilize multiple interaction domains for proper targeting, and thus a broad diversity of distinct signaling complexes may be assembled. Considering the coordination of only two such domains, as in tandem Src homology 2 (SH2) domain constructs, gives rise to a kinetic scheme that is not adequately described by simple models used routinely to interpret *in vitro* binding measurements. To analyze the interactions between tandem SH2 domains and bisphosphorylated peptides, we formulated detailed kinetic models and applied them to the phosphoinositide 3-kinase (PI3K) p85 regulatory subunit/platelet-derived growth factor (PDGF) β-receptor system. Data for this system from different *in vitro* assay platforms, including surface plasmon resonance (SPR), competition binding, and isothermal titration calorimetry (ITC), were reconciled in order to estimate the magnitude of the cooperativity characterizing the sequential binding of the high and low affinity SH2 domains (C-SH2 and N-SH2, respectively). Compared with values based on an effective volume approximation, the estimated cooperativity is three orders of magnitude lower, indicative of significant structural constraints. Homodimerization of full-length p85 was found to be an alternative mechanism for high-avidity binding to phosphorylated PDGF receptors, which would render the N-SH2 domain dispensable for receptor binding.

4.2 INTRODUCTION

Intracellular signal transduction networks, under the control of activated cell surface receptors, govern cell functional behaviors such as proliferation, migration, differentiation, and programmed cell death [1]. Proper communication between signaling proteins is generally contingent upon noncovalent, intermolecular interactions, mediated by well-conserved protein domains. A key feature of these domains is their modular nature, which has facilitated the extensive characterization of their binding affinities and specificities *in vitro*, as well as the construction of "synthetic" signaling proteins with prescribed function [2]. The prototypical and best-characterized interaction domains in signaling are the Src homology 2 (SH2) domains, which direct interactions of proteins with receptor tyrosine kinases (RTKs) and other tyrosine-phosphorylated proteins [3]. Receptors of the RTK family, which engage growth factor ligands such as platelet-derived growth factor (PDGF), are activated through ligand-binding, receptor oligomerization, and autophosphorylation on multiple intracellular residues, which then serve as a scaffold for recruitment of proteins containing SH2 and analogous domains [4, 5].

Signaling proteins typically contain three or more modular interaction domains of various types, and therefore the diversity of interactions that might take place in the cell is staggering [6]. Further complicating the problem is the avidity effect, which tends to promote the cooperative association of different domains with binding partners in the same multi-molecular complex or subcellular compartment. Other mechanisms of binding cooperativity might also depend on the modification of signaling proteins at multiple sites [7]. This context-dependent diversity of interactions is a prime example of what has been called combinatorial complexity [8]. While kinetic modeling has emerged as a powerful tool in the analysis of signal transduction networks [9-11], the very large number of potential state variables that can arise even for combinations of a handful of proteins has prohibited detailed modeling of signaling interactions. The recent development of rule-based modeling tools [12] has enabled modeling of more complex systems; in previous work, we used this approach to

analyze the function of the protein-tyrosine phosphatase Shp2 [13], demonstrating the application of rule-based modeling at the level of modular protein domains.

In this paper, we present mathematical models and analysis focused on the interactions between tandem SH2 domains derived from signal transduction proteins and peptides or proteins bearing two phospho-tyrosine binding sites. Such interactions have been characterized *in vitro* by a variety of biochemical methods [14-18], but the various types of complexes that can form between multi-valent binding partners cannot be resolved, making the measurements potentially difficult to interpret. Although dual SH2 domains are found in a number of signaling proteins, including isoforms of phospholipase C, the aforementioned Shp2, and the non-receptor tyrosine kinases Syk and ZAP70, we focus in particular on the interactions between the p85 regulatory subunit of phosphoinositide 3-kinase (PI3K) and sequences derived from the PDGF β-receptor. PI3Ks are lipid kinases that are strongly activated by PDGF receptors and by many other cell surface receptors, and they play pivotal roles in cell migration, survival, and proliferation pathways [19, 20]. The interactions of the p85 SH2 domains are critical for targeting and allosteric activation of the enzyme in cells [21-24].

Analysis of the models reconciles various published *in vitro* p85/phospho-peptide binding studies that have utilized different assay platforms, namely surface plasmon resonance (SPR) and other solid-phase binding assays, competition binding, and isothermal titration calorimetry (ITC). Thus, the consensus magnitude of the cooperativity parameter characterizing the sequential association of the two SH2 domains was evaluated and found to be orders of magnitude lower than expected based on search volume considerations. We address the implications of this apparent structural constraint in the context of PI3K recruitment and activation in cells.

4.3 METHODS

4.3.1 General modeling considerations and implementation

Our kinetic models are executed in the second-generation version of the rule-based modeling software, BioNetGen [25]. BioNetGen 2, which is freely available for download from http://bionetgen.org, uses a programming syntax that was described in detail in the supplementary material of Barua et al. [13] (also provided in Chapter 2). Graph theoretic methods are used to automatically generate a complete set of kinetic equations (ordinary differential equations in time) based on a set of user-specified rules. In this modeling framework, molecules and complexes thereof are called species, and distinct domains/motifs within the molecules are called components. Other nomenclature specific to the models presented here is as follows. The phospho-peptide has two components, Y1 and Y2, which represent phosphorylated Tyr^{751} and Tyr^{740} of the human PDGF β-receptor, respectively. The tandem SH2 construct also has two components, C-SH2 and N-SH2, corresponding to the C-terminal and more N-terminal SH2 domains of p85, respectively. Components are easily silenced in the model, by removing their corresponding rules, in order to accommodate peptides with a single phosphorylation site or p85 constructs with only one of the SH2 domains. Each of the four combinations of interactions between phospho-tyrosine and SH2 components is assigned a second-order association rate constant k_{on} and a first-order dissociation rate constant k_{off}, which characterize the reversible binding of two species to form one (Fig. 4.1 *a*). At equilibrium, it is only the ratio of these rate constants that matters, with $K_D = k_{off}/k_{on}$ given in units of molar concentration; incidentally, we used the same realistic value of $k_{on} = 1\ \mu M^{-1}s^{-1}$ for all interactions and models, and k_{off} values were specified according to the corresponding K_D. Ring closure interactions, wherein phospho-tyrosine and SH2 components within the same species associate (Fig. 4.1 *b*), are characterized by a first-order association rate constant that is the product of the corresponding k_{on} and a conversion factor χ with units of concentration [13, 26], assumed to be the same value for all ring complexes. The reverse, ring opening rate constant is given by the corresponding k_{off}. The assumption that only the forward rate constant is modified affects the binding kinetics but not

the equilibrium. The kinetic equations were integrated numerically for sufficient time to achieve steady state (10^4 seconds, typically). All model codes are available upon request.

4.3.2 Model 1: Immobilized phospho-peptide

In the simplest model, the bisphosphorylated peptide is immobilized to a surface or solid matrix, and the tandem SH2 construct binds from solution. It is assumed that the immobilized peptide is present at a sufficiently low density, such that bound complexes are comprised of only one peptide and either one or two tandem SH2 molecules. The peptide is present at an arbitrarily low concentration (10 pM was used) so that the tandem SH2 domain is far in excess, with its free concentration approximately equal to the total. Each of the peptide phosphorylation sites (Y1 or Y2), if unoccupied, may reversibly bind tandem SH2 from solution (both SH2 domains must be unoccupied) via C-SH2 or N-SH2; these 4 combinations constitute separate rules (Fig. 4.1 *a*). A peptide/SH2 complex with Y1 or Y2 unoccupied may engage in reversible ring closure transitions (4 separate rules shown in Fig. 4.1 *b*). As a result, there are 12 distinct species in this model, the 2 unbound molecules and 10 distinct peptide/SH2 complexes; the complexes are classified as Type I, II, or III depending on their structure (Fig. 4.1 *c*).

4.3.3 Model 2: Immobilized phospho-peptide with competition

This model is the same as the previous except that the system also includes soluble, bisphosphorylated peptide as a competitive inhibitor with respect to tandem SH2 binding to the surface, which allows several types of extended structures to form (Fig. 4.1 *d*). Although complexes may contain only one immobilized peptide molecule, any species containing an unoccupied SH2 domain can combine with any other having an unoccupied competitor peptide site, and thus molecular chains with more than one bisphosphorylated peptide molecule may be formed. Chains comprised of one tandem SH2 and two peptide molecules are classified as Type IV complexes, and chains comprised of four or more molecules are classified as Type V complexes. Ring structures containing four or more molecules can also form; these are classified as Type VI complexes. To simplify matters, the immobilized peptide is only mono-phosphorylated (on Y1, corresponding to $pTyr^{751}$ of PDGF β-receptor),

matching the conditions of the published experiments [17]. Thus, there are only 2 rules for C-SH2 or N-SH2 binding to the surface, 4 rules for binding of two species containing unoccupied SH2 and competitor peptide sites, and 4 rules for unimolecular ring closure involving unoccupied SH2 and competitor peptide sites. In BioNetGen 2, it is possible to set the maximum number of each molecule type in the generated species. Thus, crosslinking of immobilized sites was prohibited here by setting the maximum number of immobilized peptide molecules in a complex to 1, and the potentially infinite sizes of the chain and ring structures were truncated at a maximum number of N molecules each of the tandem SH2 and bisphosphorylated competitor peptide per complex. Values of $N = 2$, 3, and 4 were used and found to give nearly identical results. These models vary in complexity as N is increased, yielding 68, 272, and 1,075 distinct species, respectively.

4.3.4 Model 3: Solution-phase binding

In this model, both the tandem SH2 construct and bisphosphorylated peptide are in solution, as in ITC measurements. The binding rules are the same as in the immobilized phospho-peptide with competition model, except that the immobilized peptide is absent. Thus, for the same value of N as described for Model 2, there are correspondingly fewer distinct species in Model 3 (37, 145, and 629 species for $N = 2$, 3, and 4, respectively). As with Model 2, these values of N produced nearly identical results.

4.3.5 Model 4: Immobilized phospho-peptide with p85 dimerization

This model is a modification of Model 1, in which p85 has an additional domain that mediates p85 dimerization (Fig. 4.1*d*), with 6 additional rules. Two of these are for dimerization, one for when at least one of the p85 molecules binds from solution, and another for when both p85 molecules are bound to the same peptide; in the latter case, the χ value for dimerization, χ_{dimer}, is distinguished from that of 1:1 ring formation (Type II complex), called χ_{SH2}. To satisfy the principle of detailed balance, χ_{dimer} also applies to the ring closure of peptide–p85–p85 chains via either of the unoccupied SH2 domains in the second p85 molecule. The network for this model is comprised of 35 distinct species.

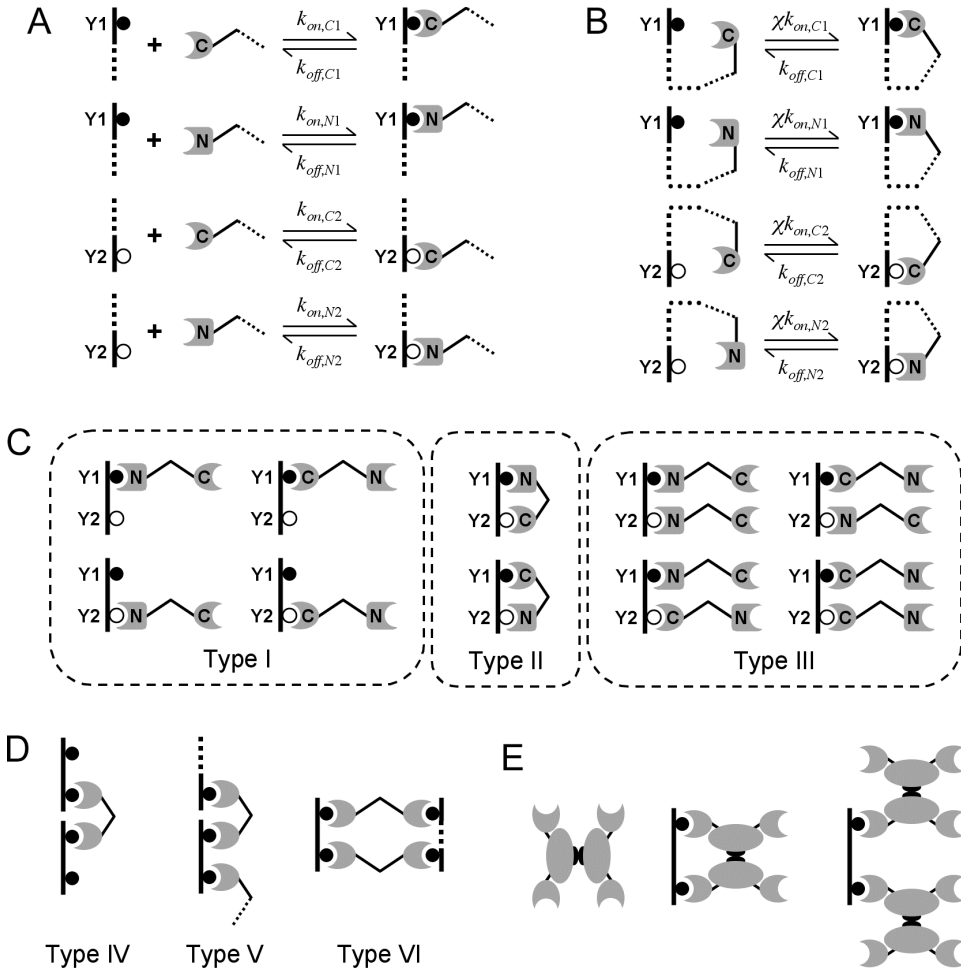


FIGURE 4.1 Rule-based model of tandem SH2 binding to bisphosphorylated peptide. (*a*) Rules for bimolecular complex formation and associated rate constants. The dashed lines indicate that the remainder of each species is unknown, potentially subject to context-dependent rules. (*b*) Ring closure transitions and associated rate constants. The cooperativity factor χ has units of concentration and applies to all such transitions. (*c*) All 10 of the distinct tandem SH2/phospho-peptide complexes containing one peptide molecule, as in the case where the peptide is immobilized at low density. Type I complexes contain one tandem SH2 domain molecule that is singly bound , Type II complexes contain one tandem SH2 domain that is doubly bound forming a ring, and Type III complexes contain two singly-bound tandem SH2 domain molecules. (*d*) Classification of chain and ring structures containing more than one peptide molecule. Type IV complexes are 1:2 chains, while Type V complexes are chains with 2:2 or higher stoichiometry. Type VI complexes are ring structures with 2:2 or higher stoichiometry. *E*. Examples of complex structures that can form when dimerization of full-length protein, such as PI3K p85, is considered.

4 RESULTS

4.4.1 Cooperativity of tandem SH2/phospho-peptide binding as a key determinant of complex avidity, stoichiometry, and equilibration time

Tandem SH2 domains, such as in the p85 regulatory subunit of PI3K, engage cognate bisphosphorylated peptides and proteins in a cooperative manner, with binding of one SH2 domain facilitating the binding of the other through a ring closure transition (Fig. 4.1*b*). This feature of the system is quantified by the parameter χ in our models, which is the effective concentration of each free binding site within the same molecular complex. If such a site were able to freely search a characteristic volume of 100 nm^3 (within a 3 nm radius), that concentration would be ~ 20 mM. A more conservative estimate would account for the flexibility of the peptide and other structural constraints within the complex [27], and hence we varied χ between 1 μM and 1 mM and evaluated its effect on the overall binding avidity and other aspects of complex formation.

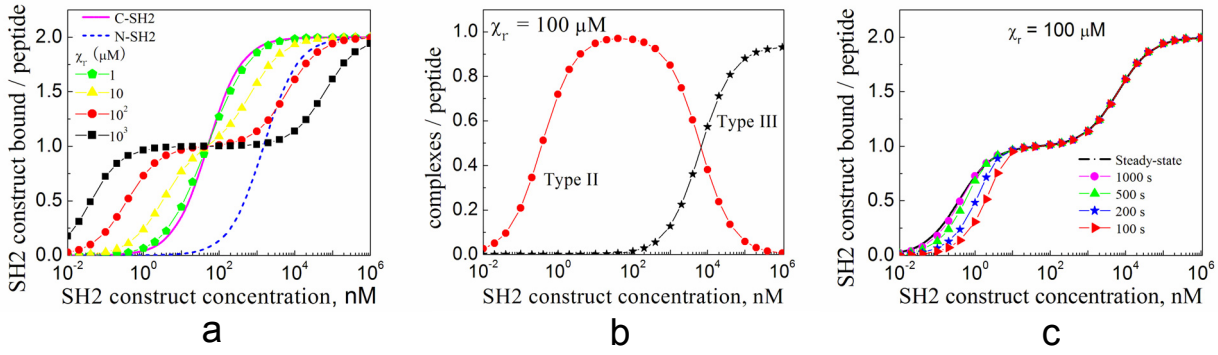


FIGURE 4.2 Binding properties of tandem SH2 constructs to immobilized, bisphosphorylated peptides. Calculations were performed using Model 1, assuming SH2 domain K_D values characteristic of PI3K p85. Constant parameter values were $k_{on,C1} = k_{on,C2} = k_{on,N1} = k_{on,N2} = 1\ \mu M^{-1}s^{-1}$, $K_{D,C1} = K_{D,C2} = 50$ nM, $K_{D,N1} = K_{D,N2} = 1.5$ μM. (*a*) Equilibrium binding isotherms. The value of χ was varied as indicated. (*b*) Structure types of complexes formed with $\chi = 100$ μM (refer to Fig. 4.1 *c*). (*c*) Tandem SH2 binding as a function of time ($\chi =$ 100 μM), with $t = 100$, 200, 500, and 1,000 seconds.

The simplest model is one in which the phospho-peptide is immobilized, such that complexes contain only one peptide molecule (Model 1) (Fig. 4.2). This scenario simulates SPR and other solid-phase binding assays and is analogous to p85 recruitment to the plasma

membrane. For simplicity, the two phospho-tyrosine sites are assumed here to be equivalent, and the C-SH2 and N-SH2 domains are assigned single-site K_D values characteristic of PI3K p85 (50 nM and 1.5 μM, respectively) [16-18, 28]. Each SH2 domain by itself exhibits the expected hyperbolic binding isotherm, with half-maximal binding at a SH2 concentration equal to its K_D and a stoichiometry of 2:1 SH2 molecules per peptide at saturation. By comparison, the binding isotherm of the tandem construct is altered relative to that of the higher affinity C-SH2, depending on the value of χ. As expected, the change is dramatic when $\chi \sim 10$ μM or greater, exceeding the K_D of the low affinity N-SH2 domain (Fig. 4.2 *a*). At tandem SH2 concentrations below the C-SH2 K_D, overall binding is enhanced because of the cooperativity of the SH2 domains in forming stable, Type II ring structures (Fig. 4.1 *c*), the effective K_D for these structures being given by

$$K_{D,eff} = \frac{1}{\chi}\left(\frac{1}{K_{D,C1}K_{D,N2}} + \frac{1}{K_{D,C2}K_{D,N1}}\right)^{-1}. \qquad (4.1)$$

Effective K_D values for p85 tandem SH2 binding to the pTyr^{740}/pTyr^{751} bisphosphorylated peptide have been reported to lie in the vicinity of 1 nM [16, 18]; for the single-site K_D values assumed here, an order-of-magnitude estimate of $\chi \sim 30$ μM is obtained. A somewhat lower estimate ($\chi \sim 10$ μM) is obtained if $K_{D,C1}$ and $K_{D,C2}$ are allowed to adopt different values spanning the range of 10-100 nM.

In contrast, at tandem SH2 concentrations above the K_D of C-SH2, overall binding is diminished because the ring structure reduces the overall stoichiometry of SH2 binding. Indeed, as the value of χ is increased, there is an apparent saturation of binding at 1:1 stoichiometry, and increasingly higher tandem SH2 concentrations are needed to shift the equilibrium from Type II rings to Type III chain structures with 2:1 stoichiometry (Fig. 4.2*b*).

Another consequence of cooperative tandem SH2 binding is slower binding kinetics (Fig. 4.2 *c*). For a simple receptor/ligand system with 1:1 binding stoichiometry, it is well known that the characteristic time constant for approaching equilibrium is the inverse of $k_{off}(1 + [L]/K_D)$, where k_{off} is the dissociation rate constant and $[L]$ is the free ligand concentration [29]. Formation of the Type II ring structure effectively increases the dwell time of the

tandem SH2 molecule on the peptide, thus reducing the overall off-rate and slowing the approach to steady state. Indeed, with the highest values of χ the $t_{1/2}$ for approaching steady state at low concentrations is greater than 5 minutes, compared with $t_{1/2} = \ln 2/k_{off} = 14$ seconds for C-SH2 alone.

4.4.2 Analysis of tandem SH2/phospho-peptide interactions in competition binding experiments establishes a lower limit on the cooperativity parameter χ

To further characterize the cooperativity of tandem SH2/phospho-peptide binding, we analyzed the data of Harpur and colleagues [17], who assessed the ability of pTyr^{740}, pTyr^{751}, and pTyr^{740}/pTyr^{751} peptides to inhibit the binding of various p85 constructs (C-SH2, tandem SH2, as well as full-length) to a SPR chip bearing pTyr^{751}; this experiment is recapitulated in

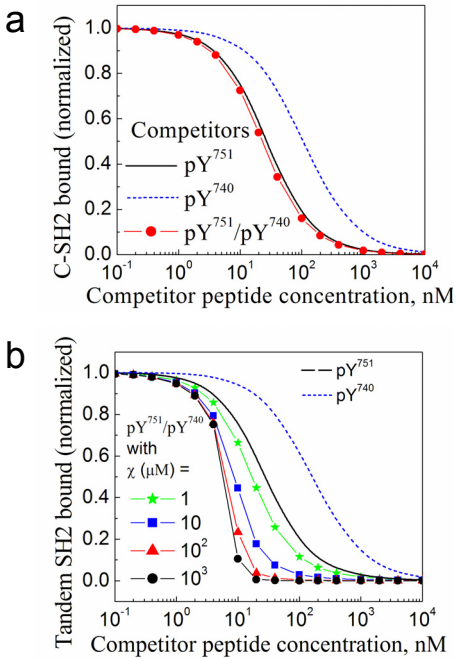


FIGURE 4.3 Evaluation of competition binding experiments. Calculations were performed using Model 2 (see Methods). (*a*) Inhibition of mono-valent p85 C-SH2 domain binding to pTyr^{751} of PDGF β-receptor by different competitor peptides as indicated. Fig. 6A of ref. [17] was recapitulated with $S_T = K_{D,C1} = 10$ nM and $K_{D,C2} = 75$ nM, where S_T is the total concentration of C-SH2. (*b*) Inhibition of p85 tandem SH2 construct by different competitor peptides as indicated. In the case of the bisphosphorylated competitor, the value of χ is varied as indicated. Fig. 6B&C of ref. [17] compare favorably with these results when $\chi >$ 10 μM.

our Model 2 (Fig. 4.3). In the relatively simple case of C-SH2 and mono-phosphorylated peptide as the competitor, the fractional occupancy of the immobilized peptide sites, assumed to be small in number compared to the SH2 molecules, is given by

$$\text{Bound fraction} = \frac{S_{free}}{K_{D,C1} + S_{free}};$$
$$S_{free} = \frac{b + \left[b^2 + 4c\right]^{1/2}}{2}; \qquad (4.2)$$
$$b = S_T - K_{D,Ci} - C_T; \quad c = K_{D,Ci} S_T .$$

Eq. 4.2 shows that the free SH2 concentration, S_{free}, depends on the total concentrations of both SH2 (S_T) and the peptide competitor (C_T) and the K_D of C-SH2 binding to the competitor site (i = 1 for pTyr^{751}, 2 for pTyr^{740}). It was assumed that S_T was chosen to yield ≈ 50% surface occupancy in the absence of competitor. Hence, good agreement with the C-SH2 inhibition data was found with $S_T = K_{D,C1}$ = 10 nM and $K_{D,C2}$ = 75 nM (Fig. 4.3 *a*), and those parameter values were kept the same in Fig. 4.3 *b*, described below. The estimate of S_T = 10 nM is corroborated elsewhere [30].

The experiments also showed that whereas each of the mono-phosphorylated competitor peptides inhibits C-SH2 and tandem SH2 binding with roughly the same potency, the bisphosphorylated competitor peptide showed enhanced potency towards tandem SH2 and full-length p85 binding, indicative of the cooperative formation of ring structures; the inhibition curve also exhibited a much steeper dose response [17]. Those results are matched nicely by the Model 2 calculations when the value of χ is much greater than the N-SH2 affinity (χ ~ 10 μM or greater; Fig. 4.3 *b*). Interestingly, the shape of the inhibition curve is not attributable to the multi-valent nature of the competitor binding, but rather to the near stoichiometric avidity of the interaction. When the binding avidity is arbitrarily high, the fractional occupancy of immobilized peptide (Eq. 4.2) is closely approximated by taking S_{free} ≈ $S_T - C_T$, or S_{free} ≈ 0 when the competitor is in excess ($C_T > S_T$), which produces the characteristic steepness of the inhibition curve. A dramatic reduction of S_T, to a value well below the effective K_D of ring formation (Eq. 4.1), eliminates this feature (results not shown); however, doing so would reduce the fractional occupancy on the surface in the absence of competitor, perhaps to an unacceptably low level for SPR detection.

4.4.3 Analysis of tandem SH2/phospho-peptide interactions in ITC measurements establishes an upper limit on the cooperativity parameter χ

ITC experiments provide information about molecular interactions through measurements of heat liberated upon serial injections of one solution into another [30]. O'Brien and colleagues performed such experiments with full-length p85, injecting increasing amounts of bisphosphorylated $pTyr^{740}/pTyr^{751}$ peptide into the calorimeter; the net energy change required to maintain the system at constant temperature with each injection was plotted as a function of the increasing molar ratio of peptide/p85 [18]. Two distinct changes in the heat per injection were observed, one starting at a molar ratio ≈ 0.5 and another, more dramatic reduction induced at a molar ratio ≈ 1.0; at a molar ratio of 2.0, the heat released was near zero, indicating saturation of the SH2 domains. Based on those molar ratios, a conceptual model was proposed in which the predominant complex at lower peptide concentrations is the 2:1 chain (Type III complex), whereas a 1:1 complex (depicted as a Type II ring) dominates for molar ratios approaching 1.0 [18].

Our calculated results (Model 3), which allow us to resolve the various types of complexes, shed additional light on those conclusions and provide further evidence for the magnitude of χ (Fig. 4.4). Based on a concentration of 10 μM p85 in the calorimeter initially and given that 1.5 nmol peptide was introduced per injection, achieving a molar ratio of 2.0 after 16 injections of 15 μL each [18], the total concentrations of p85 and peptide after each injection were determined. Thus, the total peptide concentration increases from 1.2 μM after the first injection up to 16.7 μM at the end; the p85 is diluted in the process, with a final concentration of 8.3 μM. Based on those concentrations, and using the same default K_D values from Fig. 4.2, we determined the net changes in the amounts of complexes after each injection. Changes in these amounts are related to changes in enthalpy and thus the amount of energy required to maintain constant temperature after each injection. For the sake of simplicity, we adopt a thermodynamic model in which the enthalpy change (ΔH) is a weighted sum of the numbers of bonds formed with C-SH2 and N-SH2 (n_{CSH2} and n_{NSH2}, respectively), regardless of the structures of the complexes formed:

$$\Delta H = \Delta H_{CSH2} n_{CSH2} + \Delta H_{NSH2} n_{NSH2}. \qquad (4.3)$$

This is equivalent to assuming that the induced proximity effect that distinguishes ring closure from chain extension equilibria is attributed to a difference in conformational entropy.At lower values of χ (1 and 10 μM), the calculated numbers of high-affinity C-SH2 bonds formed with each injection show the characteristic plateau at low molar ratios, thereafter yielding to formation of N-SH2 interactions, whereas for higher values of χ (100 μM and 1 mM), the plateau is absent (Fig. 4.4 *a*). Using Eq. 4.3 to calculate the heat release per injection and varying the ratio of specific enthalpies ($\Delta H_{NSH2}/\Delta H_{CSH2}$), only χ ~ 10-30 μM correctly recapitulates the experimentally observed hump in the heat per injection at molar ratios between 0.5 and 1.0 (Fig. 4.4 *b* and Fig. B3, Appendix B). Allowing the individual K_D values to adopt various values within the reported ranges, yielded similar results (Fig. B3, Appendix B).

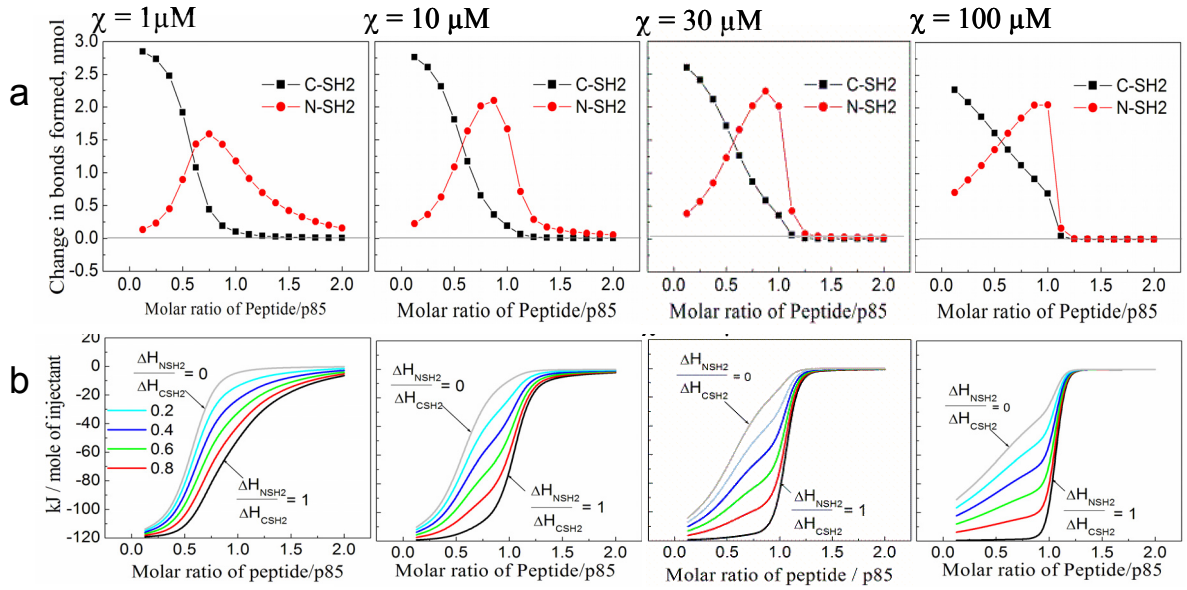


FIGURE 4.4 Evaluation of isothermal titration calorimetry (ITC) experiments. Calculations were performed using the solution-phase binding model (Model 3). In ITC experiments performed by O'Brien et al. [18], 16 aliquots of bisphosphorylated pY^{751}/pY^{740} peptide were added sequentially to a fixed amount of p85 in solution, eventually reaching a molar ratio of 2.0 peptide molecules per p85. Total p85 and peptide concentrations were determined as described in the main text, and single-site K_D values are as assumed in Fig. 2. The value of χ is given above each set of panels. (*a*)Net change in the numbers of C-SH2 and N-SH2 bonds formed with each injection of peptide. (*b*) Hypothetical enthalpy change with each injection of peptide, with ΔH_{CSH2} = –60 kJ/mol for C-SH2 bonds, and various ratios of $\Delta H_{NSH2}/\Delta H_{CSH2}$; the curves are with $\Delta H_{NSH2}/\Delta H_{CSH2}$ = 0.0 (grey) 0.2 (cyan) , 0.4 (blue), 0.6 (green), 0.8 (red), and 1.0 (black).

Further analysis of the complexes formed revealed that when the molar ratio is between 0.5 and 1, χ = 1 μM produces a mixture of chain structures of Type I and ring structures of Types II and IV, whereas χ = 10 μM leads predominantly to the formation of ring structures. In both cases, there is a shift to 1:2 (Type IV) chains as the molar ratio is increased above 1.0. In contrast, with χ = 100 μM, the shift from Type III chains to Type II rings proceeds steadily for molar ratios up to 1.0, and with χ = 1 mM, the Type II ring structure dominates throughout the hypothetical ITC run (Fig. B4, Appendix B). Taking the results of this and the previous sections together, it is suggested that the order-of-magnitude value of χ, characterizing the cooperativity of both SH2 domains of PI3K p85 engaging bisphosphorylated peptides derived from PDGF β-receptor, is 10 μM.

4.4.4 To what extent can dimerization of p85 stabilize p85 binding to bisphosphorylated peptide?

It has been shown that purified PI3K p85 dimerizes *in vitro* via a Src homology 3 (SH3) domain/proline-rich sequence interaction, estimated to be of micromolar affinity, perhaps aided by a second, lower affinity interaction [17, 31, 32]. These domains are not present in truncated, p85-derived tandem SH2 constructs, but in the context of full-length p85 we were curious as to how p85 dimerization might affect p85 interactions with the bisphosphorylated $pTyr^{740}$/$pTyr^{751}$ peptide (Fig. 4.5). In the corresponding model, Model 4, 85 dimerization is treated as a single interaction with K_D = 1 μM in solution. As in Model 1, the phospho-peptide is assumed to be immobilized at low density. Here, the structural constraints governing the formation of Type II rings are distinguished from those governing ring formation via dimerization of p85 molecules attached to the same peptide chain, characterized by distinct values of χ, χ_{SH2} and χ_{dimer}, respectively. The principle of detailed balance dictates that χ_{dimer} also applies to the cyclization of ring structures via one of the two unoccupied SH2 domains of a p85 molecule already dimerized with another, peptide-bound p85 molecule.

Assuming a value of χ_{SH2} = 10 μM, consistent with the analysis in the previous sections, the calculations show that p85 dimerization can improve binding avidity at low nanomolar concentrations, but only when rings involving dimers are not subject to significant constraints; χ_{dimer} must be in the millimolar range (Fig. 4.5 *a*). With low values of χ_{dimer}, comparable to χ_{SH2}, the binding avidity is not substantially enhanced beyond what is achieved through Type II ring formation (compare with Fig. 4.2 *a*). Ring structures with dimerized p85 molecules are found in proportion to the free p85 concentration squared, which is manifested in the steepness of the binding isotherm at low p85 concentrations. At p85 concentrations that are far in excess of the dimerization K_D, complexes with stoichiometry approaching 4:1 (two p85 dimers per peptide) are found. This model was also adapted to examine the binding of a p85 variant with the N-SH2 domain deleted (Fig. 4.5 *b*). Here, the only ring structure that can form is the 2:1 complex with the p85 molecules dimerized. Comparing the isotherm with that of wild-type p85 in Fig. 4.5 *a*, it is apparent that such rings are the predominant structure at low concentrations of p85 if χ_{dimer} is sufficiently high. Under those conditions, the N-SH2 domain is dispensable for binding to the bisphosphorylated motif.

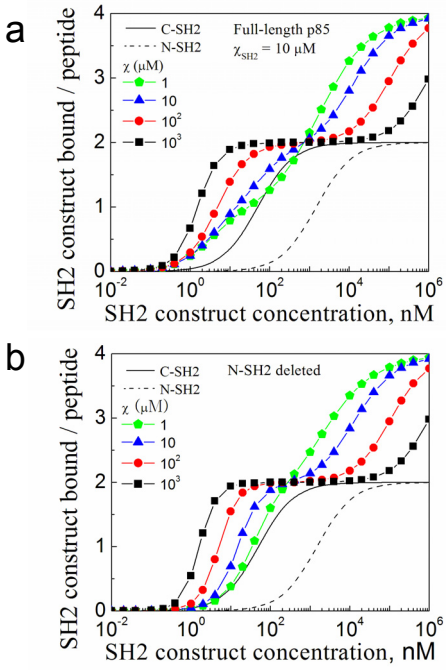


FIGURE 4.5 Effect of p85 dimerization on binding to immobilized, bisphosphorylated peptides. Calculations were performed using the immobilized phospho-peptide with p85 dimerization model (Model 4). Self-association of the dimerization domain in solution is characterized by a dissociation constant $K_{D,dimer}$ = 1 μM, and formation of Type II rings (Fig. 4.1*c*) is distinguished from other ring closure transitions by assignment of distinct χ values, χ_{SH2} and χ_{dimer}, respectively (as explained in the text). K_D values for the C-SH2 and N-SH2 domains are the same as in Figs. 4.2 and 4.4. (*a*) Equilibrium binding isotherm of full-length p85, relative to C-SH2 or N-SH2 alone, with χ_{SH2} = 10 μM and χ_{dimer} varied as indicated. (*b*) Same as A, but with N-SH2

4.5 DISCUSSION

In vitro measurements using purified components are predicated on the notion that they are indicative of interactions in cells, and they afford obvious advantages. However, when an interaction involves more than one discrete step, and especially when complexes of varying stoichiometry can form, the interpretation of the measurements can be challenging and perhaps misleading. Reconciling data obtained using different assay designs and platforms only adds to that challenge; here, we used kinetic, rule-based models to accomplish this goal. Interactions between the tandem SH2 domains of PI3K p85 regulatory subunit and its bisphosphorylated binding site in PDGF β-receptor were analyzed in detail, and the cooperativity of the SH2 domains in forming a high-avidity ring complex was evaluated in terms of the concentration factor, χ. Analysis of SPR and ITC measurements, which differ with respect to peptide configuration (immobilized versus soluble) and species concentrations (nanomolar versus micromolar), yielded a consistent order-of-magnitude estimate of $\chi \sim 10$ μM. Significantly lower values do not yield the effective K_D values reported for tandem SH2 binding to $pTyr^{740}$/$pTyr^{751}$ [16, 18], nor do they give the extent of inhibition observed in competition binding assays [17]. Significantly higher values promote ring formation even when one of the components is in micromolar excess, in clear disagreement with ITC measurements [18].

The estimate of χ obtained for p85 tandem SH2 binding is three orders of magnitude lower than the value anticipated based on simple search volume considerations, indicating significant structural constraints. Consistent with this conclusion, a worm-like chain model of peptide binding shows that consideration of the peptide flexibility alone can yield χ values in the low micromolar range [27]. However, in experiments in which the length of the peptide spacer sequence between the pTyr sites was varied, peptide stimulation of PI3K kinase activity *in vitro* was apparently able to tolerate a reduction of the spacing from 11 to 6 residues [33]; based on this assertion, the worm-like chain model produces a significantly higher estimate of $\chi \sim 10$ mM [27]. It seems clear that factors other than peptide flexibility,

such as the conformational dynamics of the tandem SH2 construct [34] and the nature of the peptide residues flanking the pTyr sites [35] and other peptide residues, must contribute to the structural constraints of the interaction.

Although still sufficient to enhance the binding of the tandem SH2 construct, the cooperativity of bisphosphorylated peptide recognition is deemed to be relatively weak, which has a number of implications for PI3K interactions with PDGF receptors in cells. Absent from experiments with receptor-derived peptides are the activities of the receptor tyrosine kinase and non-receptor tyrosine kinases that associate with activated receptors. That is significant because p85 is tyrosine-phosphorylated in cells stimulated with PDGF, on a site that engages the N-SH2 domain [36, 37]. Although it is presently unclear whether or not that interaction is intramolecular (which might lend further insights into the conformational dynamics of the p85 SH2 domains), what is clear is that the role of the interaction is to relieve the autoinhibition of PI3K catalytic activity. In our previous analysis of Shp2, which is regulated by its N-SH2 domain in a similar fashion, it was shown that Shp2 phosphorylation and intramolecular N-SH2 binding gives rise to a receptor-binding avidity that lies between two extremes; one of these is the case in which phosphorylation does not occur, and the other is the case where the N-SH2 is completely buffered from receptor binding [13]. The modest value of χ for p85/receptor binding might represent a compromise between a need for selective recognition of activated PDGF receptors, as PI3K is recruited from the cytosol, and a need for displacement of N-SH2 from the receptor after p85 is phosphorylated.

If the above is true, then the implication is that the N-SH2 domain does not contribute to PDGF receptor binding in cells to the same extent as it does to binding of bisphosphorylated peptide *in vitro*. Indeed, it has been shown that removing the N-SH2 domain of p85 does not alter its binding to PDGF receptors, but intriguingly, neither does mutation of the phosphorylation site [36, 37], suggesting that the N-SH2 domain is dispensable for receptor binding. PI3K and PDGF receptor bind extraordinarily tightly [38], and PI3K signaling stimulated by PDGF is saturated at much lower concentrations than is PDGF receptor phosphorylation [39, 40], suggesting that interactions other than C-SH2

binding to the receptor are required to stabilize the complex. Our model calculations show that p85 dimerization, whether by SH3 domain/proline-rich sequence or N-SH2/phospho-tyrosine interactions, could carry out this function, in a manner that renders the N-SH2 domain dispensable. In the context of PDGF receptor binding in cells, it is important to consider also the dimerization of PDGF receptors. This configuration might contribute parallel binding sites for the C-SH2 domains of two dimerized p85 molecules, such that the complex is less structurally constrained than in the case of binding to a single peptide or receptor molecule. Of course, interactions of the SH3 and proline-rich motifs with other molecules [32], not to mention those of the catalytic domain with substrate and possibly other binding partners, could also contribute to the stability of PI3K recruitment in cells.

4.6 REFERENCES

1. Hunter, T., Signaling– 2000 and beyond. Cell, 2000. 100: p. 113-127.

2. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A., Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu. Rev. Biochem., 2006. 75: p. 655-680.

3. Songyang, Z., and Cantley, L. C., Recognition and specificity in protein tyrosine kinase-mediated signalling. Trends Biochem. Sci., 1995. 20: p. 470-475.

4. van der Geer, P., T. Hunter, and R.A. Lindberg, Receptor protein-tyrosine kinases and their signal transduction pathways. Annu. Rev. Cell Biol., 1994. 10: p. 251-337.

5. Schlessinger, J., Cell signaling by receptor tyrosine kinases. Cell, 2000. 103(2): p. 211-225.

6. Pawson, T., Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. Cell, 2004. 116: p. 191-203.

7. Lenz, P. and P.S. Swain, An entropic mechanism to generate highly cooperative and specific binding from protein phosphorylations. Curr. Biol., 2006. 16: p. 2150-2155.

8. Barua, D., J. R. Faeder, and J. M. Haugh, The complexity of complexes in signal transduction. Biotechnol. Bioeng., 2003. 84: p. 783-794.

9. Eungdamrong, N.J. and R. Iyengar, Computational approaches for modeling regulatory cellular networks. Trends Cell Biol., 2004. 14: p. 661-669.

10. Kholodenko, B.N., Cell-signalling dynamics in time and space. Nat. Rev. Mol. Cell. Biol., 2006. 7: p. 165-176.

11. Haugh, J.M. and M.C. Weiger, Modeling intracellular signal transduction processes, in Chemical Biology: From Small Molecules to Systems Biology and Drug Design, G. Wess, Editor. 2007, Wiley-VCH.

12. Hlavacek, W. S., J. R. Faeder, M. L. Blinov, R. G. Posner, M. Hucka, and W. Fontana, Rules for modeling signal-transduction systems. Sci. STKE, 2006. 344: p. re6.

13. Barua, D., J.R. Faeder, and J.M. Haugh, Structure-based kinetic models of modular signaling protein function: focus on Shp2. Biophys. J., 2007. 92: p. 2290-2300.

14. Pluskey, S., T. J. Wandless, C. T. Walsh, and S. E. Shoelson, Potent stimulation of SHPTPs phosphatase activity by simultaneous occupancy of both SH2 domains. J. Biol. Chem., 1995. 270: p. 2897-2900.

15. Eck, M. J., S. Pluskey, T. Trub, S. C. Harrison, and S. E. Shoelson, Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. Nature, 1996. 379: p. 277-280.

16. Ottinger, E.A., M.C. Botfield, and S.E. Shoelson, Tandem SH2 domains confer high specificity in tyrosine kinase signaling. J. Biol. Chem., 1998. 273: p. 729-735.

17. Harpur, A. G., M. J. Layton, P. Das, M. J. Bottomley, G. Panayotou, P. C. Driscoll, and M. D. Waterfield, Intermolecular interactions of the p85a regulatory subunit of phosphatidylinositol 3-kinase. J. Biol. Chem., 1999. 274(18): p. 12323-12332.

18. O'Brien, R., P. Rugman, D. Renzoni, M. Layton, R. Handa, K. Hilyard, M. D. Waterfield, P. C. Driscoll, and J. E. Ladbury, Alternative modes of binding of proteins with tandem SH2 domains. Protein Sci., 2000. 9: p. 570-579.

19. Vanhaesebroeck, B. and M.D. Waterfield, Signaling by distinct classes of phosphoinositide 3-kinases. Exp. Cell Res., 1999. 253: p. 239-254.

20. Rameh, L.E. and L.C. Cantley, The role of phosphoinositide 3-kinase lipid products in cell function. J. Biol. Chem., 1999. 274(13): p. 8347-8350.

21. McGlade, C. J., C. Ellis, M. Reedijk, D. Anderson, G. Mbamalu, A. D. Reith, G. Panayotou, P. End, A. Bernstein, and A. Kazlauskas, SH2 domains of the p85a subunit of phosphatidylinositol 3-kinase regulate binding to growth factor receptors. Mol. Cell. Biol., 1992. 12: p. 991-997.

22. Panayotou, G. B., B. Bax, I. Gout, M. Federwisch, B. Wroblowski, R. Dhand, M.J. Fry, T.L.Blundell, A. Wollmer, M.D. Waterfield, Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2 domain with a PDGF receptor phosphorylation site - structural features and analysis of conformational changes. EMBO J., 1992. 11(12): p. 4261-4272.

23. Carpenter, C., K. Auger, M. Chanudhuri, M. Yoakim, B. Schaffhausen, S. Shoelson, and L. Cantley, Phosphoinositide 3-kinase is activated by phosphopeptides that bind to the SH2 domains of the 85-kDa subunit. J. Biol. Chem., 1993. 268: p. 9478-9483.

24. Shoelson, S. E., M. Sivaraja, K.P. Williams, P. Hu, J. Schlessinger and M.A. Weiss, Specific phosphopeptide binding regulates a conformational change in the PI 3-kinase SH2 domain associated with enzyme activation. EMBO J., 1993. 12(2): p. 795-802.

25. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek, BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics, 2004. 20: p. 3289-3291.

26. Haugh, J.M., I.C. Schneider, and J.M. Lewis, On the cross-regulation of protein tyrosine phosphatases and receptor tyrosine kinases in intracellular signaling. J. Theor. Biol., 2004. 230: p. 119-132.

27. Zhou, H.-X., Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen. J. Mol. Biol., 2003. 329: p. 1-8.

28. Piccione, E., R. D. Case, S. M. Domchek, P. Hu, M. Chaudhuri, J. M. Backer, J. Schlessinger, and S. E. Shoelson, Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide SH2 domain binding. Biochemistry, 1993. 32(13): p. 3197-3202.

29. Lauffenburger, D.A. and J.L. Linderman, Receptors: Models for Binding, Trafficking, and Signaling. 1993, New York: Oxford University Press.

30. Ladbury, J.E. and B.Z. Chowdhry, Sensing the heat: The application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. Chemistry & Biology, 1996. 3(10): p. 791-801.

31. Kapeller, R., K. Prasad, O. Janssen, W. Hou, B. Schaffhausen, C. Rudd, and L. Cantley, Identification of two SH3-binding motifs in the regulatory subunit of phosphatidylinositol 3-kinase. J. Biol. Chem., 1994. 269: p. 1927-1933.

32. Layton, M. J., A. G. Harpur, G. Panayotou, P. I. H. Bastiaens, and M. D. Waterfield, Binding of a diphosphotyrosine-containing peptide that mimics activated platelet-derived growth factor receptor b induces oligomerization of phosphatidylinositol 3-kinase. J. Biol. Chem., 1998. 273: p. 33379-33385.

33 Herbst, J. J., G. Andrews, L. Contillo, L. Lamphere, J. Gardner, G. E. Lienhard, and E. M. Gibbs, Potent Activation of Phosphatidylinositol 3'-Kinase by Simple Phosphotyrosine Peptides Derived from Insulin Receptor Substrate 1 Containing Two YMXM Motifs for binding SH2 Domains. Biochemistry, 1994. 33:9376-9381.

34. Siegal, G., B. Davis, S. M. Kristensen, A. Sankar, J. Linacre, R. C. Stein, G. Panayotou, M. D. Waterfield, and P. C. Driscoll, Solution structure of the C-terminal SH2 domain of the p85a regulatory subunit of phosphoinositide 3-kinase. J. Mol. Biol., 1998. 276: p. 461-478.

35. Songyang, Z., and L. C. Cantley, SH2 domains recognize specific phosphopeptide sequences. Cell, 1993. 72: p. 767-778.

36. Kavanaugh, W. M., A. Klippel, J. A. Escobedo, and L. T. Williams, Modification of the 85-kilodalton subunit of phosphatidylinositol-3 kinase in platelet-derived growth factor-stimulated cells. Mol. Cell. Biol., 1992. 12: p. 3415-3424.

37. von Willebrand, M., S. Williams, M. Saxena, J. Gilman, P. Tailor, T. Jascur, G. P., Modification of phosphatidylinositol 3-kinase SH2 domain binding properties by Abl- or Lck-mediated tyrosine phosphorylation at Tyr-688. J. Biol. Chem., 1998. 273: p. 3994-4000.

38. Kazlauskas, A. and J.A. Cooper, Phosphorylation of the PDGF receptor b-subunit creates a tight binding site for phosphatidylinositol-3 kinase. EMBO J., 1990. 9: p. 3279-3286.

39. Park, C.S., I.C. Schneider, and J.M. Haugh, Kinetic analysis of platelet-derived growth factor receptor/phosphoinositide 3-kinase/Akt signaling in fibroblasts. J. Biol. Chem., 2003. 278: p. 37064-37072.

40. Kaur, H., C.S. Park, J.M. Lewis, and J.M. Haugh, Quantitative model of Ras/phosphoinositide 3-kinase signalling cross-talk based on co-operative molecular assembly. Biochem. J., 2006. 393(1): p. 235-243.

CHAPTER 5

Quantitative model and analysis of growth hormone receptor/Jak2 signaling and the role of the SH2-Bβ adaptor

Adapted from completed manuscript (Dipak Barua, James R. Faeder, and Jason M. Haugh)

5.1 ABSTRACT

Most cell surface receptors for growth factors and cytokines dimerize in order to mediate signal transduction. For many such receptors, especially among those that govern proliferation and differentiation of hematopoietic cells, the Janus kinase (Jak) family of non-receptor tyrosine kinases are recruited in pairs and juxtaposed by dimerized receptor complexes in order to activate one another by trans-phosphorylation. Building on a rule-based kinetic modeling approach that considers the concerted interactions and combinatorial complexity of modular protein domain interactions, we examine in detail the mechanism of Jak activation and in particular the positive role of the SH2-B adaptor, focusing on the growth hormone (GH) receptor/Jak2 system. Along with the Src homology 2 (SH2) domain of SH2-B, which directly engages Jak2, it is shown that the domain responsible for SH2-B self-dimerization is critically important for activation of Jak2 in cells. Specifically, analysis of the model reveals that SH2-B dimerization enhances Jak2 autophosphorylation in conjunction with, not in addition to, Jak2 recruitment by receptor dimers. Further analysis indicates that membrane localization of SH2-B, perhaps via its pleckstrin homology (PH) domain, broadens its potency to lower intracellular concentrations.

5.2 INTRODUCTION

Growth hormone (GH) is a therapeutically important cytokine that modulates an array of cellular processes, including metabolism, proliferation, and survival [1]. Its pleiotropic effects are mediated by cell surface complexes comprised of one GH and two receptor molecules, which form as a consequence of the ordered binding of the bivalent GH ligand, a process that is understood in exquisite mechanistic detail [2]. In previous work, a model was formulated that accounted for the binding, dimerization, and trafficking of GH receptor [3]. The model successfully reconciles the potencies of certain GH mutants and monoclonal antibodies as receptor agonists or antagonists, but a notable simplification is that it relates cell proliferation in a phenomenological way to the number of dimerized GH receptors at steady state, without regard for specific intracellular signaling processes.

An essential step in GH receptor-mediated signal transduction is the activation of a non-receptor tyrosine kinase, Janus kinase (Jak) 2 [4]. Jak2 binds to a membrane-proximal, proline-rich region of growth hormone receptor [5,6], and ligand-induced receptor dimerization juxtaposes bound Jak2 to facilitate transphosphorylation of both Jak2 and the receptor [4]. Phosphorylation of Jak2 further activates the enzyme, and receptor phosphorylation sites foster recruitment of the signal transducer and activator of transcription (STAT) variants STAT3 and STAT5b, which are phosphorylated by Jak2 [7]. Given the central role of Jak2 in GH receptor signaling, it is not surprising that its function is modulated by other proteins. A prominent negative regulator is suppressor of cytokine signaling (SOCS)-1, which binds phosphorylated Tyr^{1007} in the activation loop of Jak2 and elicits degradation of the kinase [8,9]. The ubiquitously expressed adaptor protein SH2-Bβ also binds Jak2 but enhances its function [10-13]. The core structure of SH2-Bβ contains an N-terminal dimerization domain (DD), a pleckstrin homology (PH) domain, and a C-terminal Src homology-2 (SH2) domain. Among the multiple Jak2 sites phosphorylated in response to GH stimulation, Tyr^{813} is specifically recognized by the SH2-Bβ SH2 domain [14]. SH2-B also dimerizes by homotypic association of the DD, which has led to a conceptual model in which SH2-Bβ facilitates Jak2 autophosphorylation through formation of a heterotetrameric

Jak2-(SH2-Bβ)$_2$-Jak2 complex [15]. In support of this mechanism, purified SH2-Bβ enhances Jak2 phosphorylation in solution with a biphasic dose response, consistent with saturation of Jak2 at high SH2-Bβ concentrations to form dead-end Jak2-(SH2-Bβ)$_2$ complexes; in the same study, it was further shown that either the SH2 domain or DD expressed alone can antagonize GH-stimulated Jak2 and STAT5b phosphorylation in cells [15]. There is also evidence to the contrary, as the SH2 domain of SH2-Bβ was sufficient to activate Jak2 in a different context; according to this alternative mechanism, the biphasic dependence of Jak2 autophosphorylation on SH2-Bβ concentration might be attributed to a second, inhibitory interaction involving the PH domain [16,17]. Although the PH domain has not yet been characterized fully, it also has a speculated role in targeting SH2-Bβ to the plasma membrane, based on the established interactions of other PH domains with specific phosphoinositide lipids. Clearly, the two proposed mechanisms of SH2-Bβ function highlighted here present opposing views regarding the importance of DD dimerization.

In this work, we apply computational modeling to critically analyze the role of SH2-Bβ in GH receptor signaling. The model accounts for GH/GH receptor dynamics and Jak2/GH receptor, SH2/Jak2, DD/DD, and PH/lipid interactions in cells. As demonstrated in our previous domain-based models of Shp2 [18] and phosphoinositide 3-kinase regulatory subunit [19], this small number of interactions can produce thousands of distinct molecular species, and we manage this combinatorial complexity using the rule-based modeling approach [20]. Our results suggest that SH2-Bβ dimerization, although probably not a significant Jak2 association mechanism in solution, plays an important role in assembling Jak2-(GH receptor)$_2$-Jak2 complexes in the cellular context.

5.3 METHODS

5.3.1 Base model of GH/GH receptor dynamics

Where applicable, we build upon the previous model of GH/GH receptor interactions and trafficking [3] and use the same parameter values for wild-type human GH. Briefly, the GH ligand concentration [L] is fixed and is an input variable to the model, and unbound GH receptors (R) are present at a level of 2×10^3 molecules/cell initially. Receptor expression is determined by the ratio of the synthesis rate [V_s = 10 (#/cell)/min] and basal turnover rate constant (k_t = 0.005 min^{-1}). Ligand-receptor complexes (C) form with site 1 forward rate constant k_{f1} = 0.1 $nM^{-1}min^{-1}$ and reverse rate constant k_{r1} = 0.15 min^{-1} and are subject to basal turnover. Receptor dimers (D), which are competent for signaling, form from C and R with site 2 forward rate constant $k_{x2} = 2.42 \times 10^{-3}$ $(\#/cell)^{-1}min^{-1}$ and reverse rate constant k_{-x2} = 0.016 min^{-1}, and they can also dissociate via the site 1 linkage with rate constant 1.5×10^{-3} min^{-1} (as noted previously, setting this rate equal to zero does not affect the results for wild-type human GH), leaving the ligand to dissociate rapidly via the unstable site 2 linkage. Dimers are endocytosed and degraded at an enhanced rate, with rate constant k_e = 0.1 min^{-1}. Secondary effects of Jak2 and SH2-Bβ interactions on GH/GH receptor dynamics are discussed below.

5.3.2 Intracellular interactions: general considerations

Our models are based on mass-action kinetics, with bimolecular (association of two species) and unimolecular (dissociation or change in state of a complex) transitions. For all bimolecular interactions where one or both of the species is in the cytosol, the association rate constant k_{on} was assigned a typical value of 0.06 $nM^{-1}min^{-1}$ (or 1.0 $\mu M^{-1}s^{-1}$), and the dissociation rate constant k_{off} is calculated from $k_{off} = k_{on} K_D$, where K_D is the specified equilibrium dissociation constant. The total intracellular concentrations of Jak2, SH2-Bβ, and phosphoinositide (J_{Tot}, S_{Tot}, and P_{Tot}, respectively) are conserved and are specified alternatively in units of molar concentration or molecules/cell; these units are interconverted by assuming a volume of 0.52 pL, equivalent to that of a sphere with 5 μm radius.

5.3.3 Jak2 phosphorylation

Jak2 binds receptors, regardless of their ligand-bound status and the phosphorylation status of Jak2, with a K_D defined as $K_{D,RJ}$. The model considers phosphorylation of two Jak2 tyrosine sites, Y1 and Y2, corresponding to Tyr^{813} and Tyr^{1007}, which are responsible for SH2-Bβ association and stimulated activation of Jak2 kinase activity, respectively. Consistent with the current understanding of GH receptor activation, Jak2 can be phosphorylated on Y1 and Y2 only when two Jak2 molecules are associated with the same complex (receptor or/and SH2-Bβ mediated). Once Y2 is phosphorylated, the catalytic efficiency of that Jak2 molecule increases substantially. Accordingly, we model Jak2 phosphorylation as a pseudo-first order process, and once Y2 of the Jak2 molecule acting as the enzyme is phosphorylated, its phosphorylation rate constant towards both Y1 and Y2 of the other Jak2 molecule increases from 6 min^{-1} (0.1 s^{-1}) to 60 min^{-1} (1 s^{-1}). Jak2 dephosphorylation is also modeled as a pseudo-first order process, with a rate constant of 6 min^{-1} for both Y1 and Y2; phosphorylated Y1 that is bound to SH2-Bβ is protected from dephosphorylation.

5.3.4 Interactions involving SH2-Bβ

SH2-Bβ participates in as many as three interactions, with K_D values defined as follows: its SH2 domain binds to Jak2 molecules with Y1 phosphorylated ($K_{D,JS}$), its DD dimerizes ($K_{D,SS}$), and its PH domain binds phosphoinositides ($K_{D,PS}$).

The introduction of SH2-Bβ in the system gives rise to interactions in the plane of the membrane or within a multi-molecular complex, and these occur at accelerated rates in the forward direction as compared to the situation where one or both of the interacting species is in the cytosol. Dissociation of such a linkage is assumed to occur with the same rate constant as when one or both of the dissociating components is/are released into the cytosol.

Interactions between two membrane-associated species arise as a consequence of SH2-Bβ binding to phosphoinositide lipids (*PS*) or to receptor-bound Jak2 (*RJS*, with or without ligand), which can subsequently form complexes such as PS_2P, *RJSP*, RJS_2JR, etc. To simplify the model in a manner that satisfies detailed balance, interactions in the

membrane are assigned a forward rate constant that is calculated as $\chi_m k_{on}$, with k_{on} = 0.06 $nM^{-1}min^{-1}$ = 1.91 x 10^{-4} $(\#/cell)^{-1}min^{-1}$ and χ_m defined as a common, dimensionless enhancement factor; as considered in previous signal transduction models [21,22], its value is based on a confinement layer (reduced volume) with 10 nm thickness at the membrane, yielding χ_m = [(5 μm)/3(10 nm)](10^3 nm/μm) = 167. The corresponding dissociation rate constant is assumed to be the same as for release of one or both species to the cytoplasm; this assumption could relaxed if diffusion limitations were to be considered.

Interactions within a complex (ring closure) include the association of two SH2-Bβ molecules with dangling DDs, as in the species *SJ*(*RLR*)*JS*, or association of SH2-Bβ and Jak2 in the *J*(*RLR*)JS_2 complex, for example. Ring closure is a unimolecular transition with forward rate constant calculated as $\chi_r k_{on}$, where χ_r is the effective concentration of an unbound site within the complex, assumed to be the same for all such interactions (the notation is from [22], referring to interactions within a receptor complex). A conservative value of χ_r = 100 μM was used (see [19] for a detailed discussion). Ring closure also affects GH binding, because of the ability of the JS_2J heterotetramer to dimerize receptors without ligand present. Thus, the model accounts for closure of species such as $LRJS_2JR$ via the GH (site 2)/GH receptor linkage; because GH-induced receptor dimerization normally occurs in the plane of the membrane, the association rate constant for this ring closure transition is calculated as $(\chi_r/\chi_m)k_{x2}$.

To avoid the formation of potentially infinite chains at the membrane, which would occur if GH/GH receptor dimers were clustered via JS_2J linkages (which would be a rare occurrence if accounted for), the model is constrained so that complexes may contain no more than 2 receptor molecules. All complexes containing 2 receptors, whether they contain ligand or not, are considered receptor dimers and are subject to enhanced endocytosis, with rate constant k_e = 0.1 min^{-1}. Internalized receptors cannot associate with Jak2; any Jak2 and SH2-Bβ in complex with a receptor when it is internalized (whether endocytosed by the induced or basal turnover pathway) dissociate at the normal rate.

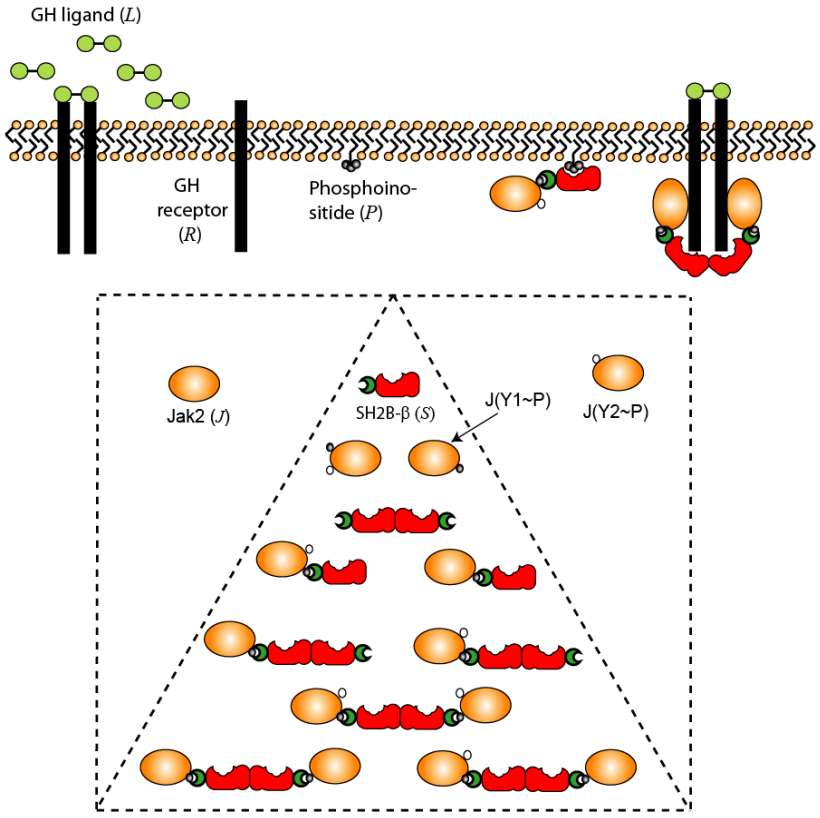


FIGURE 5.1 Molecular species and interactions considered in our models. Three models of increasing complexity were formulated and analyzed, as described under *Methods*. The In Vitro Model considers the enhancement of Jak2 (*J*) autophosphorylation by SH2-Bβ (*S*) in solution and includes 11 species (dashed triangle). Two Jak2 phosphorylation sites are considered: Y1, which when phosphorylated (Y1~P) engages SH2-Bβ, and Y2, which when phosphorylated (Y2~P) enhances the kinase activity. The model considers the best-case scenario where Y1 is constitutively (or rapidly) phosphorylated. The Simplified Cellular Model includes GH ligand (*L*) and GH receptor (*R*) and also considers Jak2 species with Y1 dephosphorylated in the cytosol (dashed square). In this model, Jak2 binds constitutively to receptors, but binding of two Jak2 molecules in the same complex is required for Jak2 autophosphorylation. Finally, the Extended Cellular Model additionally considers phosphoinositide (*P*) lipids, which mediate localization of SH2-Bβ to the plasma membrane.

5.3.5 Specific model cases and rule-based model implementation

The interactions considered in this work are summarized in Fig. 5.1. Our simplest model is the so-called In Vitro Model, which contains only Jak2 and SH2-Bβ molecules, and therefore the largest complex in this model is the heterotetramer, JS_2J. It considers the best-case scenario where all Y1 sites are pre-phosphorylated and thus generates only 11 species (state variables) (Fig. 5.1). The dephosphorylation reactions are turned off in the In Vitro

Model, because phosphatases are not present. The Simplified Cellular Model considers all of the interactions except those with phosphoinositides, generating 470 species (5,033 reactions). The Extended Cellular Model adds the influence of phosphoinositides and generates 2,561 distinct species (41,233 reactions). In variations of this model, we also considered the influence of a SH2-Bβ mutant lacking one or two of its domains, acting as a dominant negative, alongside the wild-type SH2-Bβ species; these yielded even more species and reactions, according to the complexity of the dominant negative construct considered: SH2, 2,849 species; DD, 3,154 species; PH-SH2, 3,152 species; DD-PH, 3,821 species.

Our rule-based model was developed using the software program BioNetGen2, which is freely available through http://bionetgen.org. As discussed in detail elsewhere [18], the user defines the biochemical network in terms of molecules, their interaction domains, and context-dependent rules for association/dissociation or covalent modification. Based on those rules, an exhaustive search is performed to automatically generate all possible species (combinations of interactions and modification states) and their corresponding conservation equations (differential equations in time), which are numerically integrated using a standard stiff solver up to time = 10^3 min, by which time the system was confirmed to have reached steady state. For the In Vitro Model, a time of 10 min was used, corresponding to the experimental conditions.

5.4 RESULTS

5.4.1 Jak2-SH2-Bβ heterotetramerization is an inefficient mechanism for promoting Jak2 autophosphorylation *in vitro*

Nishi et al. [15] purified Jak2 and SH2-Bβ and showed that SH2-Bβ enhances Jak2 autophosphorylation in solution. They obtained results with 14 pM Jak2 and SH2-Bβ concentrations in the range of 0.01-100 nM, which were incubated along with excess ATP for 10 minutes at 25°C in a total volume of 150 μL. The greatest change in Jak2 phosphorylation was seen as the SH2-Bβ concentration increased from 0.1 to 1 nM, and the effect of SH2-Bβ decreased at higher concentrations [15]. We recapitulated those conditions in our In Vitro Model, with the affinities of the SH2-Bβ (SH2)/Jak2 and SH2-Bβ dimerization (DD/DD) interactions varied systematically (Fig. 5.2). The SH2 domain affinity, characterized by $K_{D,JS}$, was assigned values in the range of 1-100 nM, which are at the low end of K_D values (high affinity) measured for single SH2 domains [23,24]. Indeed, although the K_D of the interaction between full-length SH2-Bβ and Jak2 is not known, the isolated SH2 domain binds to a Jak2-derived phospho-peptide with K_D = 550 nM [25]. For DD dimerization, we considered an even wider range of $K_{D,SS}$ values, from 0.1 nM to 10 μM. Because there are no phosphatases present, the dephosphorylation reactions are turned off in the In Vitro Model, and as a best-case scenario, we assume that the SH2-Bβ binding site of Jak2 (Tyr^{813}, or Y1) is pre-phosphorylated. In this context, phosphorylation of the Jak2 activation site (Tyr^{1007}, or Y2) is the readout of the model and could potentially represent multiple phosphorylation sites.

The results show that, if Jak2 autophosphorylation were to proceed by the proposed heterotetramer (JS_2J) formation mechanism, the extent of phosphorylation is at most ~ 0.3%, or < 0.01 fmol, of Jak2 (Fig. 5.2 *a–c*). Analysis of the model indicates that the rate of phosphorylation is limited by the rate of exchange between phosphorylated and unphosphorylated Jak2 in the heterotetrameric complex, which is affected by the values of the dissociation constants $k_{off,JS}$ and $k_{off,SS}$; these are determined from the assumed K_D values and typical values of the association rate constants, $k_{on,JS} = k_{on,SS} = 0.06$ $nM^{-1}min^{-1}$ (1 $\mu M^{-1}s^{-}$

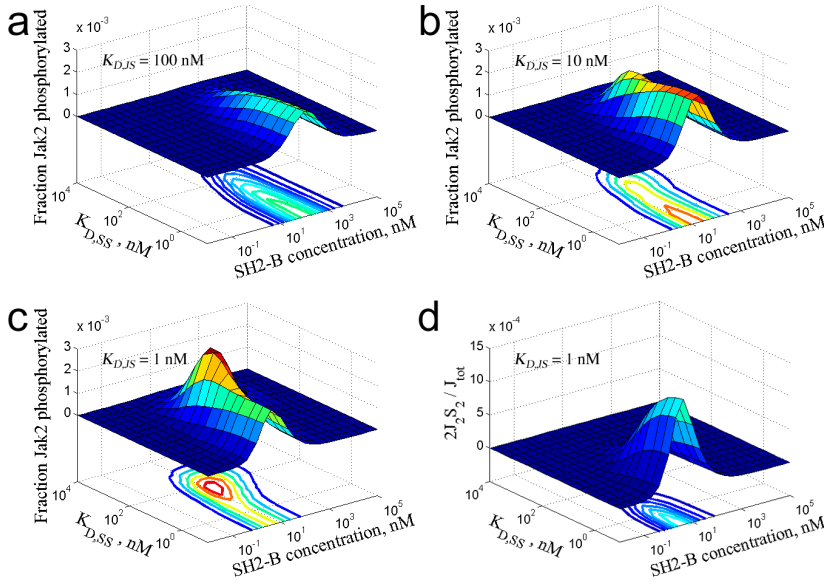


FIGURE 5.2 Critical analysis of the SH2-Bβ-mediated Jak2 autophosphorylation mechanism *in vitro*. (*a–c*) Surface and contour plots of Jak2 autophosphorylation (Y2~P) for varying concentrations and dimerization K_D values of SH2-Bβ, and with three different K_D values of Jak2/SH2-Bβ binding: (*a*) $K_{D,JS}$ = 100 nM; (*b*) $K_{D,JS}$ = 10 nM; (*c*) $K_{D,JS}$ = 1 nM. See text for a description of the model assumptions, following [15]. (*d*) Surface and contour plot of heterotetramer (JS_2J) concentration for $K_{D,JS}$ = 1 nM.

[1]). Thus, as the K_D values are decreased (higher affinity), there is a trade-off between the enhanced formation of the heterotetramer (shown in Fig. 5.2 *d* for the extreme value of $K_{D,JS}$ = 1 nM) and the reduced frequency of exchange, explaining why similar levels of Jak2 phosphorylation are predicted over multiple decades of K_D values.

These results are difficult to reconcile with the experimental observations, for the following reasons. First, to produce optimal phosphorylation at SH2-Bβ concentrations of ~ 1-10 nM, extremely high-affinity interactions are required for both the SH2 domain and DD of SH2-Bβ (K_D values ~ 1 nM). Second and more critically, the predicted amount of phosphorylated Jak2 is simply too low. Even if it were 10-fold higher, as by assuming k_{on} = 10 $\mu M^{-1}s^{-1}$ (quite high for protein-protein interactions), the amount would be well below the limit of detection for immunoblotting.

The *in vitro* role of SH2-Bβ dimerization is that much more difficult to reconcile if we relax the assumption that the SH2-Bβ binding site (Y1) is pre-phosphorylated. Indeed, an

alternative model was considered that includes SH2-Bβ-independent Jak2 dimerization and phosphorylation of Y1 as a prerequisite for SH2-Bβ binding, and we found that very high concentrations of SH2-Bβ (>> 100 nM) are needed to enhance Jak2 phosphorylation, even when the K_D values are arbitrarily high; even then, the magnitude of the enhancement is quite small (results not shown). In that model, SH2-Bβ must associate rapidly with Jak2 dimers that happen to have catalyzed the phosphorylation of Y1 on both Jak2 molecules, but not of the activating site, Y2; Y2 phosphorylation on either Jak2 molecule leads to rapid phosphorylation of available sites, in which case SH2-Bβ binding has no bearing on the Jak2 phosphorylation status of that complex. With a total Jak2 concentration of 14 pM, the overall concentration of monomeric Jak2 with Y1 phosphorylated never achieves an appreciable concentration for dimerization of Jak2/SH2-Bβ complexes in solution.

5.4.2 SH2-Bβ dimerization significantly enhances Jak2 autophosphorylation in the cellular context by coordinating Jak2/GH receptor binding

Whereas it seems unlikely that SH2-Bβ-mediated heterotetramers could form to a significant extent in solution to explain the *in vitro* data obtained by Nishi et al., Jak2 kinase activity is normally associated with cytokine receptor signaling at the plasma membrane *in vivo*. Using our Simplified Cellular Model, we quantified activated (receptor-bound and Y2-phosphorylated) Jak2 stimulated by varying doses of GH at steady state, relative to the number of cell-surface GH receptors in the absence of GH (Fig. 5.3); as explained previously [3], maximal GH receptor activation is accompanied by significant downregulation from the surface, so a relative value of ~ 0.05 by this measure is the maximum. The Simplified Cellular Model does not allow for membrane localization of SH2-Bβ through its PH domain.

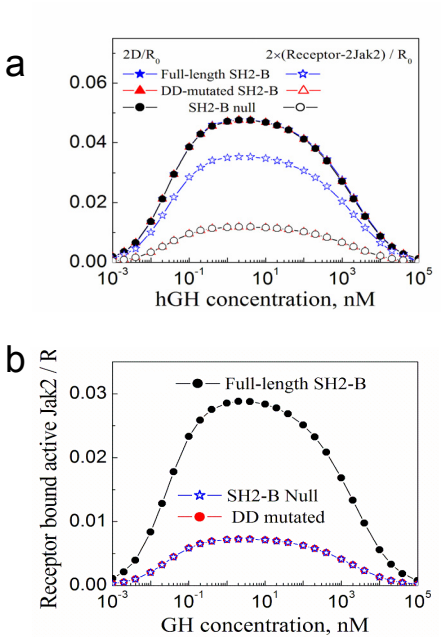


FIGURE 5.3 SH2-Bβ significantly enhances GH receptor-mediated Jak2 autophosphorylation *in vivo*. Calculations were performed using the Simplified Cellular Model, with equal Jak2 and SH2-Bβ concentrations and K_D values (J_{Tot} = S_{Tot} = $K_{D,JS}$ = $K_{D,SS}$ = 100 nM)Fig. In the SH2-B null case, S_{Tot} = 0, and in the DD-mutated SH2-B case, $K_{D,SS}$ = infinity; these two cases are functionally equivalent. (*a*) SH2-Bβ does not affect GH dose-dependent receptor-dimerization but mediates ~ 3-fold improvement in pair-wise recruitment of Jak2 to receptors (the number of Jak2 molecules engaged in receptor-Jak2 complexes containing two Jak2). (*b*) Accordingly, SH2-Bβ enhances Jak2 autophosphorylation (site Y2) by roughly 3-fold.

In the absence of SH2-Bβ, or (equivalently) with SH2-Bβ lacking the DD, the Jak2/receptor binding may be estimated in a straightforward manner. For the parameter values assumed, with total Jak2 expression in excess over receptors and equal to the K_D of Jak2/receptor binding, roughly half of the dimerized receptors are bound with Jak2, and so roughly 1/4 of the receptor dimers have two Jak2 molecules bound at steady state. Compared with that level, the presence of SH2-Bβ (with the reasonable assumption that S_{Tot} = J_{Tot} = $K_{D,JS}$ = $K_{D,SS}$) increases by ~ 3-fold the number of receptor dimers with two Jak2 bound (Fig. 5.3 *a*) and, accordingly, the number of Jak2 molecules with Y2 phosphorylated (Fig. 5.3 *b*). We reason that it does so by forming stable, seven-member "macro-complexes" containing GH, two receptor, two Jak2, and two dimerized SH2-Bβ molecules, as depicted in Fig. 5.1. To further characterize this hypothetical mechanism, the intracellular concentration and dimerization affinity of SH2-Bβ were varied for a constant GH concentration of 10 nM (Fig. 5.4). Although a broad range of SH2-Bβ concentrations was tested in order to evaluate the full spectrum of behaviors, it is noted that the endogenous SH2-Bβ expression level is not expected to be above the nanomolar range. Given a constant Jak2/SH2-Bβ affinity ($K_{D,JS}$ =

100 nM), the SH2-Bβ concentration should be of a similar magnitude or somewhat higher for near maximal enhancement of Jak2 phosphorylation; extremely high SH2-Bβ concentrations, similar in magnitude to χ_r (100 μM) are needed to antagonize the formation of the stable macro-complex, leading instead to formation of less stable, nine-member $S_2J(RLR)JS_2$ complexes (Fig. 5.4 *a*). Analysis of the GH receptor/Jak2 complexes formed reveals that, as expected, SH2-Bβ stabilizes complexes with two Jak2 molecules while increasing the total Jak2 recruitment only modestly (Fig. 5.4 *b*).

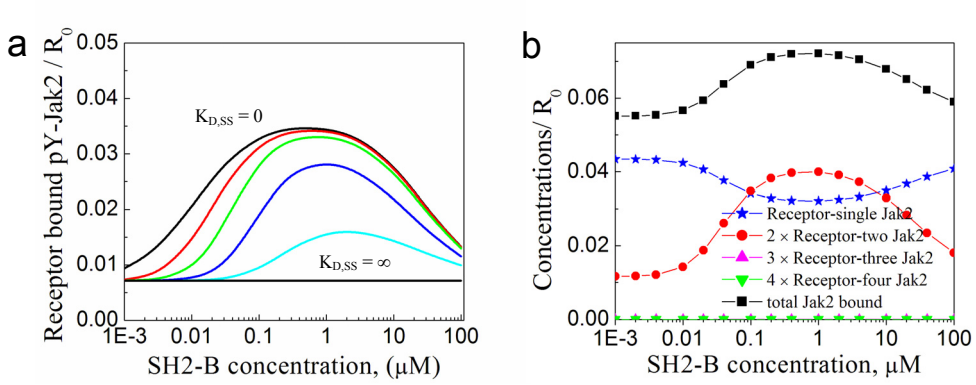


FIGURE 5.4 SH2-Bβ dimerization coordinates the formation of macro-complexes containing two Jak2 molecules bound to GH-dimerized receptors. Steady-state calculations were performed using the Simplified Cellular Model and the same parameter values as in Fig. 5.3, except with 10 nM GH stimulation and varying SH2-Bβ concentration. *a*) Receptor-bound, phosphorylated Jak2 (Y2~P), for various values of the SH2-Bβ dimerization affinity. The extreme cases of $K_{D,SS}$ equal to zero and infinity correspond to irreversible and no dimerization, respectively; intermediate $K_{D,SS}$ values are 10 nM (red), 100 nM (green), 1 μM (blue), and 10 μM (cyan). *b*) Analysis of receptor/Jak2 complexes, with $K_{D,SS}$ = 100 nM. SH2-Bβ dimerization coordinates the binding of two Jak2 molecules to dimerized receptors, while affecting overall receptor/Jak2 binding only modestly. Complexes containing more than two Jak2 molecules (e.g., $J(RLR)JS_2J$) are rare.

5.4.3 Membrane localization of SH2-Bβ via its PH domain broadens Jak2 activation potency, but SH2-Bβ dimerization is still essential

We next considered the role of the SH2-Bβ PH domain, which is thought to mediate binding with phosphoinositides and thus plasma membrane localization [26], in our Extended Cellular Model (Fig. 5.5). Based on physical principles, membrane localization increases the rate of association between complexes containing receptor or/and phosphoinositide molecules by roughly two orders of magnitude, enhancing the binding of SH2-Bβ with

receptor-bound Jak2. In fact, we find that the addition of the PH domain interaction broadens the efficacy of SH2-Bβ-mediated Jak2 activation down to low nanomolar SH2-Bβ concentrations, well below the assumed K_D of the Jak2/SH2-Bβ interaction in solution (Fig. 5.5 *a*). As in the Simplified Cellular Model, this enhancement is not accompanied by dramatic gains in overall Jak2/receptor binding (Fig. 5.5 *b*). Membrane-localization of SH2-Bβ facilitates binding to receptor-bound Jak2 and SH2-Bβ dimerization, and therefore it

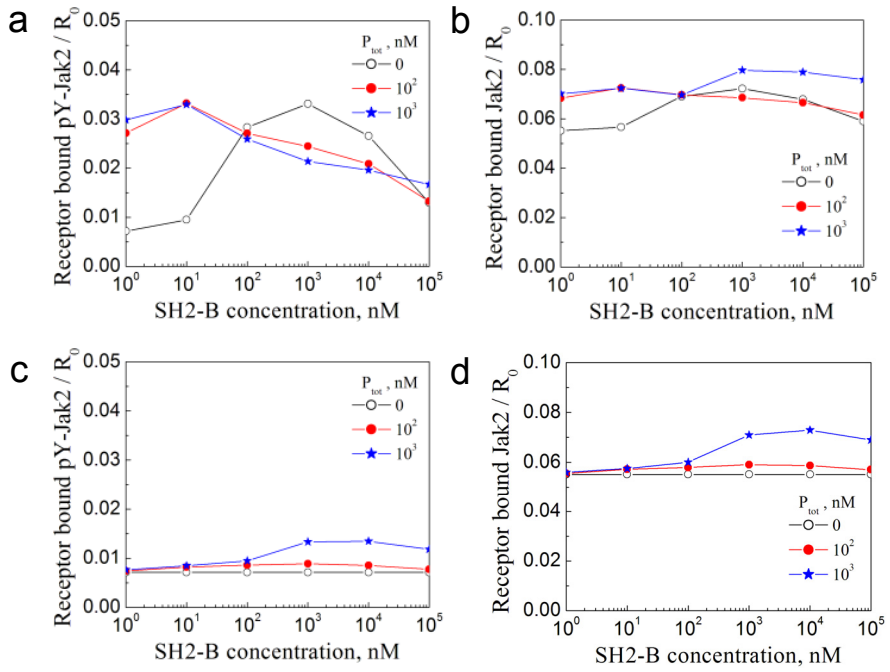


FIGURE 5.5 Membrane localization and dimerization of SH2-Bβ synergize to enhance the potency of its Jak2 activation-promoting function. Steady-state calculations were performed using the Extended Cellular Model and the same parameter values as in Fig. 4 *b*. The total concentration of phosphoinositide, on a whole-cell basis, is either 0, 100 nM, or 1 μM as indicated, and its recruitment of SH2-Bβ PH domain is characterized by $K_{D,SP}$ = 100 nM. Two scenarios are considered: full-length SH2-Bβ (*a* & *b*) and SH2-Bβ with the dimerization domain absent (*c* & *d*). The calculated quantities are receptor-bound, phosphorylated Jak2 (Y2~P; *a* & *c*) and total receptor-bound Jak2 (*b* & *d*).

stabilizes signaling-competent macro-complexes at the expense of other receptor/Jak2 complexes.

To probe this mechanism further, we repeated the analysis with the DD of SH2-Bβ removed. Intuitively, one might expect that membrane localization of SH2-Bβ would drive significantly more Jak2 into complex with receptors; however, this was not the case with the DD present (Fig. 5.5 *b*), and accordingly, Jak2 autophosphorylation was not dramatically enhanced by SH2-Bβ with the DD absent, even with arbitrarily high SH2-Bβ and phosphoinositide concentrations (Fig. 5.5 *c* & *d*). Variation of the other parameters, such as the Jak2 concentration and binding affinities, did not qualitatively affect the outcome (results not shown).

Why is SH2-Bβ dimerization predicted to be so important in the cellular context? A key insight is that Jak2 must be phosphorylated on Y1, by associating with dimerized receptors, before it can bind membrane-localized SH2-Bβ. Phosphorylated Jak2 might even associate with SH2-Bβ quite readily, but the lifetime of the receptor/Jak2 interaction is not affected as a result, and the association of *JSP* complexes with free receptors is modest because this pool of Jak2 is small; once formed, the *JSP* complex is more likely to dissociate via one of its two linkages then to associate with a free receptor site, and when it does bind free receptors, it does not discriminate between dimerized and inactive receptor molecules. By comparison, SH2-Bβ dimerization specifically stabilizes Jak2 interactions with dimerized receptors.

5.4.4 Predictions regarding the potency of SH2-Bβ mutants as dominant-negative inhibitors of GH receptor signaling

To further evaluate the roles of the functional SH2-Bβ domains, we assessed the ability of different domain mutants to antagonize the function of wild-type SH2-Bβ in cells, i.e., to act as a dominant negative (Fig. 5.6). The Extended Cellular Model was used with the addition of the mutant SH2-Bβ species. The SH2 domain alone competes with wild-type for Jak2 binding and is an effective inhibitor at concentrations of at least 1 μM (for nanomolar concentrations of endogenous SH2-Bβ, as expected), which is 10-fold above the assumed value of $K_{D,JS}$ (Fig. 5.6 *a*). Inhibition by the DD alone is through dimerization with wild-type

SH2-Bβ and is somewhat less effective (Fig. 5.6 *b*), which might be attributed to the partial neutralization of the DD through homo-dimerization. The addition of the PH domain to either the SH2 domain or the DD results in membrane localization of the mutant SH2-Bβ and, accordingly, more potent disruption of receptor/Jak2/SH2-Bβ macro-complexes when it is expressed in excess compared with wild-type SH2-Bβ; comparing PH-SH2 and DD-PH, the former construct shows the more robust inhibition of SH2-B-β function (Fig. 5.6 *c* & *d*).

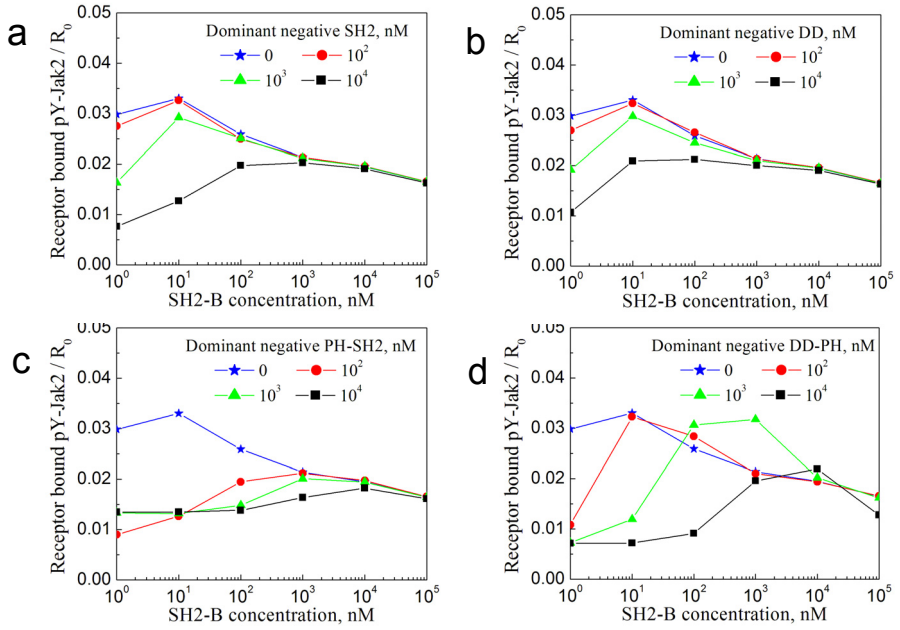


FIGURE 5.6 Potencies of SH2B-β domain mutants as dominant negatives antagonizing wild-type SH2-Bβ function. Jak2 phosphorylation was calculated using the Extended Cellular Model as in Fig. 5.5 *a*, with the same parameter values and P_{Tot} = 1 μM. To this model, we added one of the following SH2-Bβ constructs: SH2 only (*a*), DD only (*b*), PH-SH2 (*c*), and DD-PH (*d*). As indicated, the value of the overall inhibitor concentration was either 0 (no inhibition; same as Fig. 5.5 *a*), 100 nM, 1 μM, or 10 μM.

5.5 DISCUSSION

This is the third system we have studied using the rule-based modeling approach to specifically address the concerted binding of multiple, modular domains in signaling proteins. This aspect of signal transduction is a recognized source of complexity in the signal transduction field [27,28], yet it is commonly (and most often prudently) side-stepped in the formulation of mathematical models of signaling pathways. At the level of pathways and networks, we recognize and espouse that the finer molecular details, while important to consider, must be simplified (or "lumped", in the mathematical sense). The rule-based approach addresses the problem of combinatorial complexity [29], its main strength being that it allows the modeler to invoke more mechanistic or biologically plausible assumptions [30]; however, it cannot ease the burden associated with specifying a large number of model parameters, which becomes increasingly problematic at the pathway/network level. For this reason, we apply rule-based modeling to "small-scale" systems that involve only a handful of interactions yet give rise to combinations of complexes that could not readily be enumerated in the classical way. Indeed, in this work, models with as many as 3,821 differential equations were generated.

We analyzed the receptor-mediated activation of Jak2 and the role of the adaptor protein SH2-Bβ, which contains three modular domains (DD, PH, and SH2), and demonstrated how modeling can be used to evaluate the integration of domain functions as they affect receptor-mediated signaling in cells. In particular, we sought to clarify the role of SH2-Bβ dimerization. Protein homodimerization, or dimerization of structurally homologous proteins, is a ubiquitous process in molecular biology and permeates signal transduction from the receptor level (e.g., cytokine receptors, receptor tyrosine kinases) to the activation of transcription factors (e.g., STATs, Smads). Ligand-induced dimerization of the GH receptor is necessary but not sufficient for intracellular signaling, requiring also the juxtaposition of two Jak2 molecules; this theme is common to the closely related erythropoietin receptor and also multi-subunit cytokine receptors that are phosphorylated by the Jak-family kinases [31,32]. Dimerization of SH2-B isoforms, and of the closely related APS proteins, is unique

because they are considered adaptors or modulators of, not executors of, intracellular signaling.

The analysis illustrates the importance of SH2-Bβ dimerization but also sheds light on the mechanism by which it enhances Jak2 phosphorylation. Specifically, it is only effective in the context of the dimerized receptor complex, which templates the assembly of the JS_2J heterotetrameric unit. Thus, dimerized receptors and SH2-Bβ together coordinate the recruitment of two Jak2 molecules. At least in the context of our models, it is incorrect to characterize SH2-Bβ dimerization as a means of bringing two Jak2 molecules together, as might be inferred by the ability of the adaptor to enhance Jak2 autophosphorylation in solution; rather, we suggest that its role is to stabilize existing $J(RLR)J$ complexes. This is because Jak2 must already be autophosphorylated, at least on Tyr^{813}, for SH2-Bβ to bind. Accordingly, enhancing the association rates of the RJ/S or R/JS linkages, as by membrane localization of SH2-Bβ, is insufficient for significant enhancement of Jak2 phosphorylation if SH2-Bβ cannot dimerize.

This work puts forward a number of testable predictions. One concerns the mechanism by which SH2-Bβ dimerization affects Jak2 autophosphorylation, as outlined above; others consider the potential role of the SH2-Bβ PH domain (or whichever structural motif is responsible for the observed membrane localization). In a cellular context where endogenous SH2-Bβ expression is lacking or repressed, comparison of wild-type SH2-Bβ and a mutant defective in lipid binding might only show moderate differences, and in fact the mutant might outperform the wild-type adaptor if the adaptor concentration is in the high nanomolar range (as is often the case for expression plasmids; Fig. 5.5). The model results suggest that the role of membrane localization is to broaden the efficacy of SH2-Bβ to low or sub-nanomolar concentrations of the adaptor. But by the same token, we show that membrane localization of SH2-Bβ should enhance the inhibitory properties of constructs that lack either the SH2 domain or the DD, and thus the importance of the membrane localization effect might be more effectively interrogated through such inhibition experiments. To put these predictions in the proper context, it will be important to identify the sequence(s) of

SH2B-β responsible for its apparent membrane localization, whether in the PH domain or elsewhere in the molecule.

5.6 REFERENCES

1. Lanning, N.J. and C. Carter-Su. 2006. Recent advances in growth hormone signaling. Rev. Endocr. Metab. Disord., 7: 225-235.

2. Wells, J.A. 1996. Binding in the growth hormone receptor complex. Proc. Natl. Acad. Sci. U.S.A., 93: 1-6.

3. Haugh, J.M. 2004. Mathematical model of human growth hormone (hGH)-stimulated cell proliferation explains the efficacy of hGH variants as receptor agonists or antagonists. Biotechnol. Prog., 20: 1337-1344.

4. Argetsinger, L.S., G.S. Campbell, X.N. Yang, B.A. Witthuhn, O. Silvennoinen, J.N. Ihle, and C. Carter-Su. 1993. Identification of Jak2 as a growth-hormone receptor-associated tyrosine kinase. Cell, 74: 237-244.

5. van der Kuur, J.A., X.Y. Wang, L.Y. Zhang, G.S. Campbell, G. Allevato, N. Billestrup, G. Norsted, and C. Carter-Su. 1994. Domains of the growth-hormone receptor required for association and activation of Jak2 tyrosine kinase. J. Biol. Chem., 269: 21709-21717.

6. Frank, S.J., W. Yi, Y. Zhao, J.F. Goldsmith, G. Gilliland, J. Jiang, I. Sakai, and A.S. Kraft. 1995. Regions of the JAK2 tyrosine kinase required for coupling to the growth hormone receptor. J. Biol. Chem., 270: 14776-14785.

7. Smit, L.S., D.J. Meyer, N. Billestrup, G. Norstedt, J. Schwartz, and C. Carter-Su. 1996. The role of the growth hormone (GH) receptor and JAK1 and JAK2 kinases in the activation of Stats 1, 3, and 5 by GH. Mol. Endocrinol., 10: 519-533.

8. Yasukawa, H., H. Misawa, H. Sakamoto, M. Masuhara, A. Sasaki, T. Wakioka, S. Ohtsuka, T. Imaizumi, T. Matsuda, J.N. Ihle, and A. Yoshimura. 1999. The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. EMBO J., 18: 1309-1320.

9. Ungureanu, D., P. Saharinen, I. Junttila, D.J. Hilton, and O. Silvennoinen. 2002. Regulation of Jak2 through the ubiquitin-proteasome pathway involves phosphorylation of Jak2 on Y1007 and interaction with SOCS-1. Mol. Cell. Biol., 22: 3316-3326.

10. Rui, L., L. Mathews, K. Hotta, T. Gustafson, and C. Carter-Su. 1997. Identification of SH2-Bbeta as a substrate of the tyrosine kinase JAK2 involved in growth hormone signaling. Mol. Cell. Biol., 17: 6633-6644.

11. Rui, L. and C. Carter-Su. 1999. Identification of SH2-Bβ as a potent cytoplasmic activator of the tyrosine kinase Janus kinase 2. Proc. Natl. Acad. Sci. U.S.A., 96: 7172-7177.

12. Yousaf, N., Y.P. Deng, Y.H. Kang, and H. Riedel. 2001. Four PSM/SH2-B alternative splice variants and their differential roles in mitogenesis. J. Biol. Chem., 276: 40940-40948.

13. O'Brien, K.B., J.J. O'Shea, and C. Carter-Su. 2002. SH2-B family members differentially regulate JAK family tyrosine kinases. J. Biol. Chem., 277: 8673-8681.

14. Kurzer, J.H., L.S. Argetsinger, Y.-J. Zhou, J.-L.K. Kouadio, J.J. O'Shea, and C. Carter-Su. 2004. Tyrosine 813 is a site of JAK2 autophosphorylation critical for activation of JAK2 by SH2-Bβ. Mol. Cell. Biol., 24: 4557-4570.

15. Nishi, M., E.D. Werner, B.-C. Oh, J.D. Frantz, S. Dhe-Paganon, L. Hansen, J. Lee, and S.E. Shoelson. 2005. Kinase activation through dimerization by human SH2-B. Mol. Cell. Biol., 25: 2607-2621.

16. Rui, L., D.R. Gunter, J. Herrington, and C. Carter-Su. 2000. Differential binding to and regulation of JAK2 by the SH2 domain and N-terminal region of SH2-Bβ. Mol. Cell. Biol., 20: 3168-3177.

17. Kurzer, J.H., P. Saharinen, O. Silvennoinen, and C. Carter-Su. 2006. Binding of SH2-B family members within a potential negative regulatory region maintains JAK2 in an active state. Mol. Cell. Biol., 26: 6381-6394.

18. Barua, D., J.R. Faeder, and J.M. Haugh. 2007. Structure-based kinetic models of modular signaling protein function: focus on Shp2. Biophys. J., 92: 2290-2300.

19. Barua, D., J.R. Faeder, and J.M. Haugh. 2008. Computational models of tandem Src homology 2 domain interactions and application to phosphoinositide 3-kinase. J. Biol. Chem., 283: 7338-7345.

20. Hlavacek, W.S., J.R. Faeder, M.L. Blinov, R.G. Posner, M. Hucka, and W. Fontana. 2006. Rules for modeling signal-transduction systems. Sci. STKE, 344: re6.

21. Haugh, J.M. and D.A. Lauffenburger. 1997. Physical modulation of intracellular signaling processes by locational regulation. Biophys. J., 72: 2014-2031.

22. Haugh, J.M., I.C. Schneider, and J.M. Lewis. 2004. On the cross-regulation of protein tyrosine phosphatases and receptor tyrosine kinases in intracellular signaling. J. Theor. Biol., 230: 119-132.

23. Ladbury, J.E., M.A. Lemmon, M. Zhou, J. Green, M.C. Botfield, and J. Schlessinger. 1995. Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. Proc. Natl. Acad. Sci. U.S.A., 92: 3199-3203.

24. Ottinger, E.A., M.C. Botfield, and S.E. Shoelson. 1998. Tandem SH2 domains confer high specificity in tyrosine kinase signaling. J. Biol. Chem., 273: 729-735.

25. Hu, J. and S.R. Hubbard. 2006. Structural basis for phosphotyrosine recognition by the Src homology-2 domains of the adapter proteins SH2-B and APS. J. Mol. Biol., 361: 69-79.

26. Rui, L., J. Herrington, and C. Carter-Su. 1999. SH2-B, a membrane-associated adapter, is phosphorylated on multiple serines/threonines in response to nerve growth factor by kinases within the MEK/ERK cascade. J. Biol. Chem., 274: 26485-26492.

27. Pawson, T. 2004. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. Cell, 116: 191-203.

28. Bhattacharyya, R.P., A. Remenyi, B.J. Yeh, and W.A. Lim. 2006. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu. Rev. Biochem., 75: 655-680.

29. Hlavacek, W.S., J.R. Faeder, M.L. Blinov, A.S. Perelson, and B. Goldstein. 2003. The complexity of complexes in signal transduction. Biotechnol. Bioeng., 84: 783-794.

30. Blinov, M.L., J.R. Faeder, B. Goldstein, and W.S. Hlavacek. 2006. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. Biosystems, 83: 136-151.

31. Constantinescu, S.N., S. Ghaffari, and H.F. Lodish. 1999. The erythropoietin receptor: Structure, activation, and intracellular signal transduction. Trends Endocrinol. Metab., 10: 18-23.

32. Kisseleva, T., S. Bhattacharya, J. Braunstein, and C.W. Schindler. 2002. Signaling through the JAK/STAT pathway, recent advances and future challenges. Gene, 285: 1-24.

CHAPTER 6

Conclusions and future work

6.1 CONCLUSIONS AND PROSPECTS FOR FUTURE WORK

In this study, we have demonstrated the use of a rule-based algorithm in modeling domain level protein interactions in cell signaling pathways. We have successfully incorporated the micro-level details in assessing the sub-molecular functional regulations of modular proteins, and analyzed the potential intracellular implications for three important signaling enzymes (Shp2, PI3K and SH2-Bβ). In addition, we have captured the potential mechanistic picture of a partial interaction network in growth hormone (GH) signaling. Our results underscore the need for assessing the functional behaviors of individual modular signaling proteins, which has largely been disregarded in traditional modeling.

Future works now can be focused to reduce our detailed models in a systematic manner by capturing their essential performance characteristics. Simplification is necessary for larger networks that cannot be modeled explicitly by the domain-based approach. We propose to subdivide large network models into small modules/subsystems and to formulate detailed models for each subsystem independently. Subsequently we want to simplify these subsystem models and integrate them in order to evaluate the core characteristics of the network.

As an example of systematic reduction, we can consider the detailed model of PI3K – PDGF-receptor interactions that we had analyzed in chapter 4. The model quantified the apparent intracomplex cooperativity (~10 μM) of tandem PI3K binding with the bisphosphorylated sequence of PDGF receptors. At this level of cooperativity, the protein should exhibit extremely high effective affinity (lower nanomolar K_D) for receptor binding. Therefore, in a stimulated cell, activated PDGF receptors are expected to bind PI3K molecules at 1:1 stoichiometric ratio. This insight helps simplify the model greatly. If the cellular concentration of PI3K is limiting with respect to activated PDGF receptors, we can

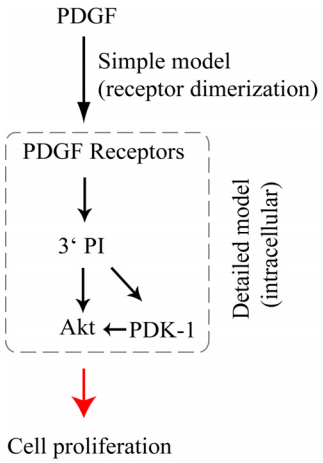


FIGURE 6.1 Reduced model for PI3K – Akt pathway. The pathway (refer to Fig. 21*a*) is simplified by omitting PI3K from the network.

assume that all PI3K molecules remain receptor-bound and kinase active. On the contrary, if receptor concentration is limiting, we can assume that all activated receptors remain saturated by PI3K and can use them directly in place of activated PI3K (Fig. 6.1). We can also simplify the cascade by assigning single site interactions between the PDGF-receptor and PI3K with an affinity corresponding to the ~ 10 μM tandem cooperativity (or we can extrapolate Fig. 4.2 *a* to obtain the net PI3K binding).

In addition to lumping the detailed interactions, we can simplify pathway models by implementing meso/macro state definitions of molecular species when permissible [1]. For instance, in the SH2-B – Jak2 pathway model, the receptor-bound and activated Jak2 molecules interact and activate the transcription factors, STATs[2, 3]. Inclusion of the STAT and further downstream network in this current model would incur massive complexity. It is to be noted that Jak2 activation and other upstream events are independent of STAT or its downstream components. Now if the current model generates 1,000 distinct molecular species that incorporate the receptor-bound, active Jak2 (say, J_i, where i = 1 - 1,000), we can classify all these species in a single group, and define the group as a single macrostate component, $\sum[J_i] = [X]$. We can subsequently allow X to interact and activate the STAT molecules. This simplification would avoid the transmission of complexity from one layer to another, still preserving the characteristics of detailed domain-based modeling.

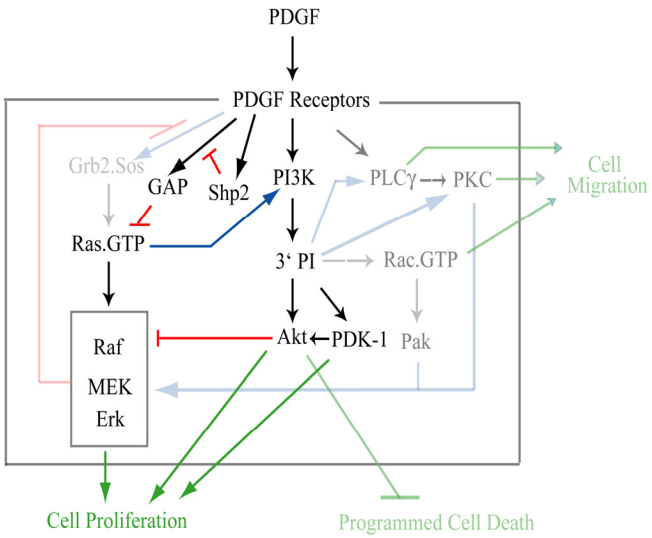


FIGURE 6.2 Erk signaling and PI3K signaling cascades. Showing the cross talk between these pathways.

Now we can think about the cross-interactions between the Erk and PI3K-Akt signaling cascades (Fig. 6.2), where Ras.GTP can affect PI3K activities. As mentioned earlier, for the PI3K – Akt signaling cascade, we can assume a single tyrosine site for the receptor and a single SH2 domain for PI3K, and allow these two sites to interact with a single-site K_D (low nanomolar) equivalent to the ~ 10 μM tandem cooperativity. This would simplify the PDGF – PI3K interaction module significantly. Secondly, we can define a macrostate (as discussed above) for all receptor-bound active PI3K, and allow it to interact with the lipid substrates at plasma membrane. In the Erk signaling cascade, we can keep our detailed Shp2 model intact, and define a macrostate for all PDGF-receptors that remain phosphorylated at Tyr^{771} (the GAP binding site of PDGF-receptors). Subsequently, we can allow this macrostate to interact and activate the GAP molecules.

In addition to model reduction, we propose to formulate simple models that will fit the results of the detailed subsystem models. The outputs of the detailed models will be treated as the experimental data in this context. We will seek to identify the quantitative relationships between the parameters of the detailed models and the estimated lumped parameters of the simple models. To obtain good fit between the detailed and simple models,

we will iteratively refine the simple models by adding more detail progressively. In parallel, the detailed models will be modified with simplifying assumptions as needed.

We believe that the rule-based modeling possesses enormous potential in guiding new experiments. We have implemented this approach in simulating experimental contexts, such as surface plasmon resonance and isothermal titration calorimetry. In addition, we have shown how this approach can be used in mimicking the protein domain mutation effects in a cell. Focus now can be directed towards exploring other biochemical methods that might also be simulated by this modeling. Many of these assays, which are often expensive and cumbersome to accomplish, could be simulated first by rule-based modeling, and this might provide the opportunity for justifying many intuitive hypotheses before carrying out actual validating experiments. Therefore, we see a bright prospect of integrating our rule-based modeling works with laboratory experimentation in near future.

6.2 REFERENCES

1. Conzelmann, H., J. Saez-Rodriguez, T. Sauter, B. Kholodenko, and E. Gilles, A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks. BMC Bioinformatics, 2006. 7(1): p. 34.

2. Feng, J., B. A. Witthuhn, T. Matsuda, F. Kohlhuber, I. M. Kerr, and J. N. Ihle, Activation of Jak2 catalytic activity requires phosphorylation of Y-1007 in the kinase activation loop. Molecular and Cellular Biology, 1997. 17(5): p. 2497-2501.

3. Schindler, C. and J.E. Darnell, Transcriptional Responses to Polypeptipe Ligands - the Jak-Stat Pathway. Annual Review of Biochemistry, 1995. 64: p. 621-651.

APPENDICES

APPENDIX A

A1 BioNetGen codes for the base model of Shp2 regulation

```
# Base model of Shp2 regulation from Barua, Faeder, and Haugh (2006).
# Copyright 2006, North Carolina State University and Los Alamos National
# Laboratory

# Concentration units are in micromolar; time units are in seconds.

version("2.0.34");

begin parameters
kopen           10
kclose          500

kon_CSH2        1
koff_CSH2       1

kon_NSH2        1
koff_NSH2       1

kkin_Y1         0.1

kon_PTP         1
koff_PTP        10
kcat_PTP        1

chi_r1          1000
chi_r2          100
chi_r3          1000
chi_r4          1000
chi_r5          100
chi_r6          100
chi_r7          100
chi_r8          1000  # Equals chi_r1*chi_r6/chi_r2
chi_r9          100   # Equals chi_r1*chi_r7/chi_r3
chi_r10         100   # Equals chi_r1*chi_r6/chi_r4
chi_r11         1000  # Equals chi_r1*chi_r7/chi_r5

R_dim         0.025 # R_tot= 2*R_dim
S_tot         0.05
end parameters

begin molecule types
R(DD,Y1~U~P,Y2~P)
S(NSH2~C~O,CSH2,PTP~C~O)
end molecule types

begin species
S(NSH2~C,CSH2,PTP~C)                            S_tot
# Pre-dimerized receptors
R(DD!1,Y1~U,Y2~P).R(DD!1,Y1~U,Y2~P)              R_dim
end species

begin reaction rules
# Intra-complex phosphorylation
```

```
R(DD!+,Y1~U) -> R(DD!+,Y1~P)  kkin_Y1

# Equilibrium between the closed form and open form of S
S(NSH2~C,PTP~C) <-> S(NSH2~O,PTP~O)  kopen,kclose

# Binding of S(CSH2) from cytosol
R(Y2~P) + S(CSH2) <-> R(Y2~P!1).S(CSH2!1)  kon_CSH2,koff_CSH2 \
exclude_reactants(2,R)

# Binding of S(NSH2~O) from cytosol
R(Y2~P) + S(NSH2~O) <-> R(Y2~P!1).S(NSH2~O!1)  kon_NSH2,koff_NSH2 \
exclude_reactants(2,R)

# Binding of S(PTP~O) from cytosol
R(Y1~P) + S(PTP~O) <-> R(Y1~P!1).S(PTP~O!1)  kon_PTP,koff_PTP \
exclude_reactants(2,R)

# Dephosphorylation of R(Y1~P)
R(Y1~P!1).S(PTP~O!1) -> R(Y1~U) + S(PTP~O)  kcat_PTP
R(Y1~P!1).S(PTP~O!1) -> R(Y1~U).S(PTP~O)  kcat_PTP

# 1 Intra-complex binding: CSH2 bound, association of NSH2 (open) with other
receptor
R(Y2~P).S(NSH2~O,CSH2!+,PTP~O) <-> \
R(Y2~P!1).S(NSH2~O!1,CSH2!+,PTP~O)  chi_r1*kon_NSH2,koff_NSH2

# 2 Intra-complex binding: CSH2 bound, association of PTP (open) with same receptor
R(Y1~P,Y2~P!1).S(NSH2~O,CSH2!1,PTP~O) <-> \
R(Y1~P!2,Y2~P!1).S(NSH2~O,CSH2!1,PTP~O!2)  chi_r2*kon_PTP,koff_PTP

# 3 Intra-complex binding: CSH2 bound, association of PTP (open) with other
receptor
R(Y1~P).R(Y2~P!1).S(NSH2~O,CSH2!1,PTP~O) <-> \
R(Y1~P!2).R(Y2~P!1).S(NSH2~O,CSH2!1,PTP~O!2)  chi_r3*kon_PTP,koff_PTP

# 4 Intra-complex binding: NSH2 bound, association of CSH2 with other receptor
R(Y2~P).S(NSH2~O!+,CSH2,PTP~O) <-> \
R(Y2~P!1).S(NSH2~O!+,CSH2!1,PTP~O)  chi_r1*kon_CSH2,koff_CSH2

# 5 Intra-complex binding: NSH2 bound, association of PTP with other receptor
R(Y1~P).R(Y2~P!1).S(NSH2~O!1,CSH2,PTP~O) <-> \
R(Y1~P!2).R(Y2~P!1).S(NSH2~O!1,CSH2,PTP~O!2)  chi_r4*kon_PTP,koff_PTP

# 6 Intracomplex binding: NSH2 bound, association of PTP with same receptor
R(Y1~P,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O) <-> \
R(Y1~P!2,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O!2)  chi_r5*kon_PTP,koff_PTP

# 7 Intra-complex binding: PTP bound, association of CSH2 with same receptor
R(Y1~P!1,Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).S(NSH2~O,CSH2!2,PTP~O!1)  chi_r2*kon_CSH2,koff_CSH2

# 8 Intra-complex binding: PTP bound, association of CSH2 with other receptor
R(Y1~P!1).R(Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1).R(Y2~P!2).S(NSH2~O,CSH2!2,PTP~O!1)  chi_r3*kon_CSH2,koff_CSH2

# 9 Intra-complex binding: PTP bound, association of NSH2 with other receptor
R(Y1~P!1).R(Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2,PTP~O!1)  chi_r4*kon_NSH2,koff_NSH2
```

```
# 10 Intra-complex binding: PTP bound, association of NSH2 with same receptor
R(Y1~P!1,Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).S(NSH2~O!2,CSH2,PTP~O!1)  chi_r5*kon_NSH2,koff_NSH2

# 11 Intra-complex binding: CSH2 & NSH2 bound, assoc. of PTP with same receptor as
CSH2
R(Y1~P,Y2~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2!1,PTP~O) <-> \
R(Y1~P!3,Y2~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2!1,PTP~O!3) \
chi_r6*kon_PTP,koff_PTP

# 12 Intra-complex binding: CSH2 & NSH2 bound, assoc. of PTP with same receptor as
NSH2
R(Y1~P,Y2~P!1).R(Y2~P!2).S(NSH2~O!1,CSH2!2,PTP~O) <-> \
R(Y1~P!3,Y2~P!1).R(Y2~P!2).S(NSH2~O!1,CSH2!2,PTP~O!3) \
chi_r7*kon_PTP,koff_PTP

# 13 Intra-complex binding: CSH2 & PTP bound to the same receptor, assoc. of NSH2
R(Y1~P!1,Y2~P!2).R(Y2~P).S(NSH2~O,CSH2!2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).R(Y2~P!3).S(NSH2~O!3,CSH2!2,PTP~O!1) \
chi_r8*kon_NSH2,koff_NSH2

# 14 Intra-complex binding: CSH2 & PTP bound to different receptors, assoc. of NSH2
R(Y2~P!1).R(Y1~P!2,Y2~P).S(NSH2~O,CSH2!1,PTP~O!2) <-> \
R(Y2~P!1).R(Y1~P!2,Y2~P!3).S(NSH2~O!3,CSH2!1,PTP~O!2) \
chi_r9*kon_NSH2,koff_NSH2

# 15 Intra-complex binding: PTP & NSH2 bound to different receptors, assoc. of CSH2
R(Y2~P!1).R(Y1~P!2,Y2~P).S(NSH2~O!1,CSH2,PTP~O!2) <-> \
R(Y2~P!1).R(Y1~P!2,Y2~P!3).S(NSH2~O!1,CSH2!3,PTP~O!2) \
chi_r10*kon_CSH2,koff_CSH2

# 16 Intra-complex binding: PTP & NSH2 bound to same receptor, assoc. of CSH2
R(Y1~P!1,Y2~P!2).R(Y2~P).S(NSH2~O!2,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).R(Y2~P!3).S(NSH2~O!2,CSH2!3,PTP~O!1) \
chi_r11*kon_CSH2,koff_CSH2

end reaction rules



begin observables
Molecules  pYR      R(Y1~P!?)
end observables

generate_network();
writeSBML();
simulate_ode({t_end=>1000,n_steps=>100,steady_state=>1,atol=>1e-10,rtol=>1e-12});
```

A2 BioNetGen codes for the extended model of Shp2 regulation

```
# Extended model of Shp2 regulation from Barua, Faeder, and Haugh (2006) that
# includes an additional regulatory mechanism activated by phosphorylation of
# tyrosine in the C-terminal tail of Shp2 that binds to the NSH2 domain.

# Copyright 2006, North Carolina State University and Los Alamos National
# Laboratory

# Concentration units are in micromolar; time units are in seconds.
```

```
version("2.0.34");

begin parameters
kopen           10
kclose          500

kon_CSH2        1
koff_CSH2       1

kon_NSH2        1
koff_NSH2       1

kkin_Y1         0.1

kon_PTP         1
koff_PTP        10
kcat_PTP        1

chi_r1          1000
chi_r2          100
chi_r3          1000
chi_r4          1000
chi_r5          100
chi_r6          100
chi_r7          100
chi_r8          1000  # Equals chi_r1*chi_r6/chi_r2
chi_r9          100   # Equals chi_r1*chi_r7/chi_r3
chi_r10         100   # Equals chi_r1*chi_r6/chi_r4
chi_r11         1000  # Equals chi_r1*chi_r7/chi_r5

R_dim        0.025 # R_tot= 2*R_dim
S_tot        0.05

# Additional parameters
kkin_Shp2       1
kdePO4          1
kon_PO4         100
koff_PO4        0.1
end parameters

begin molecule types
R(DD,Y1~U~P,Y2~P)
S(NSH2~C~O,CSH2,PTP~C~O,Y~U)
end molecule types

begin species
S(NSH2~C,CSH2,PTP~C,Y~U)              S_tot
# Pre-dimerized receptors
R(DD!1,Y1~U,Y2~P).R(DD!1,Y1~U,Y2~P)      R_dim
end species

begin reaction rules
# Intra-complex phosphorylation
R(DD!+,Y1~U) -> R(DD!+,Y1~P)  kkin_Y1

# Equilibrium between the closed form and open form of S
S(NSH2~C,PTP~C) <-> S(NSH2~O,PTP~O)  kopen,kclose
```

```
# Binding of S(CSH2) from cytosol
R(Y2~P) + S(CSH2) <-> R(Y2~P!1).S(CSH2!1)  kon_CSH2,koff_CSH2 \
exclude_reactants(2,R)

# Binding of S(NSH2~O) from cytosol
R(Y2~P) + S(NSH2~O) <-> R(Y2~P!1).S(NSH2~O!1)  kon_NSH2,koff_NSH2 \
exclude_reactants(2,R)

# Binding of S(PTP~O) from cytosol
R(Y1~P) + S(PTP~O) <-> R(Y1~P!1).S(PTP~O!1)  kon_PTP,koff_PTP \
exclude_reactants(2,R)

# Dephosphorylation of R(Y1~P)
R(Y1~P!1).S(PTP~O!1) -> R(Y1~U) + S(PTP~O)  kcat_PTP
R(Y1~P!1).S(PTP~O!1) -> R(Y1~U).S(PTP~O)  kcat_PTP

# 1 Intra-complex binding: CSH2 bound, association of NSH2 (open) with other
# receptor
R(Y2~P).S(NSH2~O,CSH2!+,PTP~O) <-> \
R(Y2~P!1).S(NSH2~O!1,CSH2!+,PTP~O)  chi_r1*kon_NSH2,koff_NSH2

# 2 Intra-complex binding: CSH2 bound, association of PTP (open) with same receptor
R(Y1~P,Y2~P!1).S(NSH2~O,CSH2!1,PTP~O) <-> \
R(Y1~P!2,Y2~P!1).S(NSH2~O,CSH2!1,PTP~O!2)  chi_r2*kon_PTP,koff_PTP

# 3 Intra-complex binding: CSH2 bound, association of PTP (open) with other
# receptor
R(Y1~P).R(Y2~P!1).S(NSH2~O,CSH2!1,PTP~O) <-> \
R(Y1~P!2).R(Y2~P!1).S(NSH2~O,CSH2!1,PTP~O!2)  chi_r3*kon_PTP,koff_PTP

# 4 Intra-complex binding: NSH2 bound, association of CSH2 with other receptor
R(Y2~P).S(NSH2~O!+,CSH2,PTP~O) <-> \
R(Y2~P!1).S(NSH2~O!+,CSH2!1,PTP~O)  chi_r1*kon_CSH2,koff_CSH2

# 5 Intra-complex binding: NSH2 bound, association of PTP with other receptor
R(Y1~P).R(Y2~P!1).S(NSH2~O!1,CSH2,PTP~O) <-> \
R(Y1~P!2).R(Y2~P!1).S(NSH2~O!1,CSH2,PTP~O!2)  chi_r4*kon_PTP,koff_PTP

# 6 Intracomplex binding: NSH2 bound, association of PTP with same receptor
R(Y1~P,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O) <-> \
R(Y1~P!2,Y2~P!1).S(NSH2~O!1,CSH2,PTP~O!2)  chi_r5*kon_PTP,koff_PTP

# 7 Intra-complex binding: PTP bound, association of CSH2 with same receptor
R(Y1~P!1,Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).S(NSH2~O,CSH2!2,PTP~O!1)  chi_r2*kon_CSH2,koff_CSH2

# 8 Intra-complex binding: PTP bound, association of CSH2 with other receptor
R(Y1~P!1).R(Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1).R(Y2~P!2).S(NSH2~O,CSH2!2,PTP~O!1)  chi_r3*kon_CSH2,koff_CSH2

# 9 Intra-complex binding: PTP bound, association of NSH2 with other receptor
R(Y1~P!1).R(Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2,PTP~O!1)  chi_r4*kon_NSH2,koff_NSH2

# 10 Intra-complex binding: PTP bound, association of NSH2 with same receptor
R(Y1~P!1,Y2~P).S(NSH2~O,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).S(NSH2~O!2,CSH2,PTP~O!1)  chi_r5*kon_NSH2,koff_NSH2

# 11 Intra-complex binding: CSH2 & NSH2 bound, assoc. of PTP with same receptor as
```

```
# CSH2
R(Y1~P,Y2~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2!1,PTP~O) <-> \
R(Y1~P!3,Y2~P!1).R(Y2~P!2).S(NSH2~O!2,CSH2!1,PTP~O!3) \
chi_r6*kon_PTP,koff_PTP

# 12 Intra-complex binding: CSH2 & NSH2 bound, assoc. of PTP with same receptor as
# NSH2
R(Y1~P,Y2~P!1).R(Y2~P!2).S(NSH2~O!1,CSH2!2,PTP~O) <-> \
R(Y1~P!3,Y2~P!1).R(Y2~P!2).S(NSH2~O!1,CSH2!2,PTP~O!3) \
chi_r7*kon_PTP,koff_PTP

# 13 Intra-complex binding: CSH2 & PTP bound to the same receptor, assoc. of NSH2
R(Y1~P!1,Y2~P!2).R(Y2~P).S(NSH2~O,CSH2!2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).R(Y2~P!3).S(NSH2~O!3,CSH2!2,PTP~O!1) \
chi_r8*kon_NSH2,koff_NSH2

# 14 Intra-complex binding: CSH2 & PTP bound to different receptors, assoc. of NSH2
R(Y2~P!1).R(Y1~P!2,Y2~P).S(NSH2~O,CSH2!1,PTP~O!2) <-> \
R(Y2~P!1).R(Y1~P!2,Y2~P!3).S(NSH2~O!3,CSH2!1,PTP~O!2) \
chi_r9*kon_NSH2,koff_NSH2

# 15 Intra-complex binding: PTP & NSH2 bound to different receptors, assoc. of CSH2
R(Y2~P!1).R(Y1~P!2,Y2~P).S(NSH2~O!1,CSH2,PTP~O!2) <-> \
R(Y2~P!1).R(Y1~P!2,Y2~P!3).S(NSH2~O!1,CSH2!3,PTP~O!2) \
chi_r10*kon_CSH2,koff_CSH2

# 16 Intra-complex binding: PTP & NSH2 bound to same receptor, assoc. of CSH2
R(Y1~P!1,Y2~P!2).R(Y2~P).S(NSH2~O!2,CSH2,PTP~O!1) <-> \
R(Y1~P!1,Y2~P!2).R(Y2~P!3).S(NSH2~O!2,CSH2!3,PTP~O!1) \
chi_r11*kon_CSH2,koff_CSH2

# ADDITIONAL RULES

# Phosphorylation of Shp2
S(Y~U) -> S(Y~P) kkin_Shp2 \
include_reactants(1,R)

# Dephosphorylation of Shp2
S(Y~P) -> S(Y~U) kdePO4

# Intramolecular binding of NSH2 (open) to the C-terminal phosphotyrosine
S(Y~P,NSH2~O) <-> S(Y~P!1,NSH2~O!1)  kon_PO4, koff_PO4

end reaction rules

begin observables
Molecules  pYR      R(Y1~P!?)
end observables

generate_network();
writeSBML();
simulate_ode({t_end=>1000,n_steps=>100,steady_state=>1,atol=>1e-10,rtol=>1e-12});
```

A3 Simplified kinetic model of Shp2 regulation

(Refer to Eq. 3.3, Chapter 3)

In the following derivation, *D* stands for deactivated (i.e., Tyr^{771} dephosphorylated) receptor-dimer, and *S* stands for unbound/cytosolic Shp2. The asterix (*) sign over *D* represents its active (Tyr^{771} phosphorylated) state. A first order rate is assumed for receptor-dimer activation. *S* binds with the active and inactive form of the receptor dimers indistinguishably. Complex D^*S is deactivated through the phosphatase activity of *S*. D^* and D^*S combined represent the total concentration of active receptor-dimers, whereas *DS* and D^*S together account for the total receptor-dimer that are in complex with S.

$$D + S \underset{k_-}{\overset{k_+}{\longleftrightarrow}} DS$$

$$D^* + S \underset{k_-}{\overset{k_+}{\longleftrightarrow}} D^*S$$

$$D \xrightarrow{k_{kin}} D^*$$

$$DS \xrightarrow{k_{kin}} D^*S$$

$$D^*S \xrightarrow{k_{PTP}} DS$$

Steady - state mass balance equations

$$-k_+[D][S] + k_-[DS] - k_{kin}[D] = 0 \qquad \text{A3.1}$$

$$-k_+[D^*][S] + k_-[D^*S] + k_{kin}[D] = 0 \qquad \text{A3.2}$$

$$k_+[D][S] - k_-[DS] - k_{kin}[DS] + k_{PTP}[D^*S] = 0 \qquad \text{A3.3}$$

$$k_+[D^*][S] - k_-[D^*S] + k_{kin}[DS] - k_{PTP}[D^*S] = 0 \qquad \text{A3.4}$$

$$D_T = [D] + [D^*] + [DS] + [D^*S] \qquad \text{A3.5}$$

$(1) \Rightarrow$

$$[DS] = \left(\frac{[S]}{K_D} + \frac{k_{kin}}{k_-} \right)[D] \qquad \text{A1.6}$$

$$[DS] = \left(\frac{b_D}{1-b_D} + \frac{k_{kin}}{k_-} \right)[D]$$

$$\therefore [D\} + [DS] = \left(1 + \frac{b_D}{1-b_D} + \frac{k_{kin}}{k_-} \right)[D]$$

$$= \frac{[D]}{1-b_D}\left(1 + (1-b_D)\frac{k_{kin}}{k_{kin}+k_{PTP}}\frac{k_{kin}+k_{PTP}}{k_-} \right)$$

$$= \frac{[D]}{1-b_D}\left(1 + (1-b_D)\phi\frac{Q}{1-Q} \right)$$

$$[D] + [DS] = \frac{1}{(1-b_D)}\frac{1}{(1-Q)}[D][1-Q+(1-b_D)\phi Q]$$

A3.6

(1) + (2) ⇒

$\frac{[S]}{K_D} = \frac{[DS]+[D^*S]}{[D\}+[D^*]}$; Where $K_D = \frac{k_-}{k_+}$

$\Rightarrow \frac{[S]}{K_D+[S]} = b_D$

$\Rightarrow \frac{[S]}{K_D} = \frac{b_D}{1-b_D}$

$$\frac{k_{kin}+k_{PTP}}{k_{kin}+k_{PTP}+k_-} = Q$$

$$\frac{k_{kin}+k_{PTP}}{k_-} = \frac{Q}{1-Q}$$

(A1.2) + (A1.4) ⇒

$$[D^*S] = \frac{k_{kin}}{k_{PTP}}([D]+[DS])$$

$$= \frac{k_{kin}}{k_{PTP}}\left([D] + \left(\frac{b_D}{1-b_D} + \frac{k_{kin}}{k_-} \right)[D] \right)$$

$$[D^*S] = \frac{k_{kin}}{k_{PTP}}\left(\frac{1}{1-b_D} + \frac{k_{kin}}{k_-} \right)[D]$$

$A1.2 \Rightarrow$

$$[D^*] = \frac{K_D}{[S]}[D^*S] + \frac{k_{kin}[D]}{k_-\frac{[S]}{K_D}}$$

$$[D^*] + [D^*S] = \left(1 + \frac{K_D}{[S]} \right)[D^*S] + \frac{k_{kin}[D]}{k_-\frac{[S]}{K_D}}$$

$$= \frac{1}{b_D}\frac{k_{kin}}{k_{PTP}}\left(\frac{1}{1-b_D} + \frac{k_{kin}}{k_-} \right)[D] + \frac{k_{kin}[D]}{k_-\left(\frac{b_D}{1-b_D} \right)}$$

$$= \frac{1}{b_D}\frac{k_{kin}}{k_{PTP}}[D]\left(\frac{1}{1-b_D} + \frac{k_{kin}}{k_-} + \frac{(1-b_D)k_{PTP}}{k_-} \right)$$

$$\frac{k_{kin}}{k_{kin}+k_{PTP}} = \phi$$

$$\frac{k_{kin}}{k_{PTP}} = \frac{\phi}{1-\phi}$$

$$\frac{k_{PTP}}{k_{kin}+k_{PTP}} = 1-\phi$$

$$= \frac{1}{b_D}\frac{k_{kin}}{k_{PTP}}[D]\left(\frac{1}{1-b_D}+\frac{k_{kin}+k_{PTP}}{k_-}+b_D\frac{k_{PTP}}{k_{kin}+k_{PTP}}\frac{k_{kin}+k_{PTP}}{k_-}\right)$$

$$= \frac{1}{b_D}\frac{\phi}{1-\phi}[D]\left(\frac{1}{1-b_D}+\frac{Q}{1-Q}+b_D(1-\phi)\frac{Q}{1-Q}\right)$$

$$[D^*]+[D^*S] = \frac{1}{b_D(1-b_D)}\frac{\phi}{(1-\phi)}\frac{1}{(1-Q)}[D]\{1-Qb_D[1+(1-\phi)(1-b_D)]\} \qquad \text{A3.7}$$

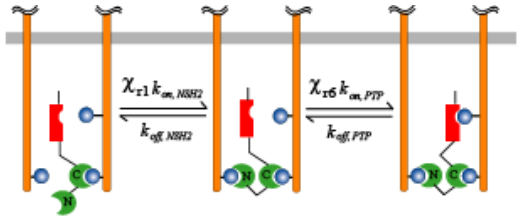
A3.7 ÷ A3.6

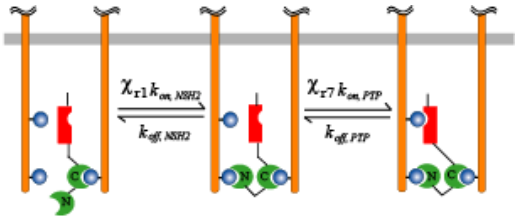
$$\frac{[D^*]+[D^*S]}{[D]+[DS]} = \frac{\phi\{1-Qb_D[1+(1-\phi)(1-b_D)]\}}{b_D(1-\phi)[1-Q+(1-b_D)\phi Q]}$$

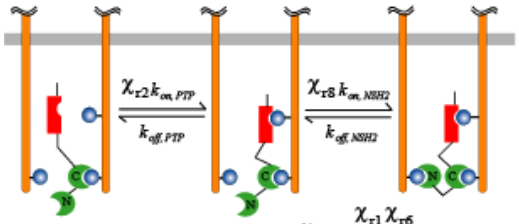
$$\frac{p_D}{1-p_D} = \frac{\phi\{1-Qb_D[1+(1-\phi)(1-b_D)]\}}{b_D(1-\phi)[1-Q+(1-b_D)\phi Q]}$$

$$p_D = \phi\left\{\frac{1-Qb_D[1+(1-\phi)(1-b_D)]}{\phi(1-b_D)+b_D(1-Q)}\right\} \qquad \text{A3.8}$$

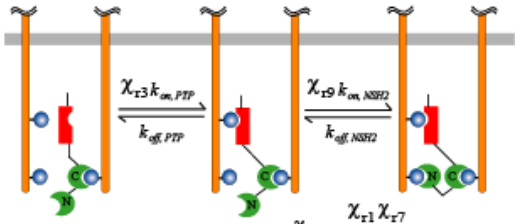
A4 Enhancement factors for cooperative intracomplex binding in Shp2 regulation models



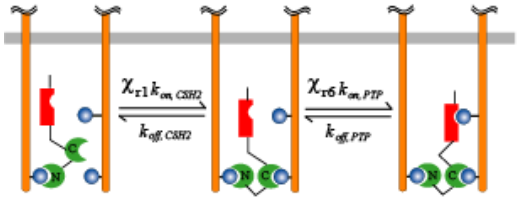


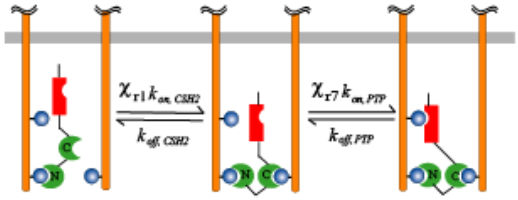


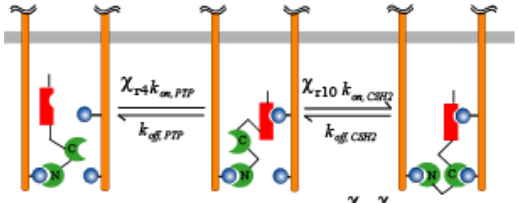
$$\chi_{r8} = \frac{\chi_{r1}\chi_{r6}}{\chi_{r2}}$$



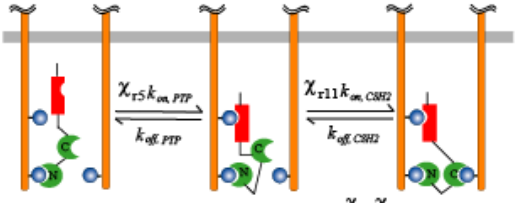
$$\chi_{r9} = \frac{\chi_{r1}\chi_{r7}}{\chi_{r3}}$$



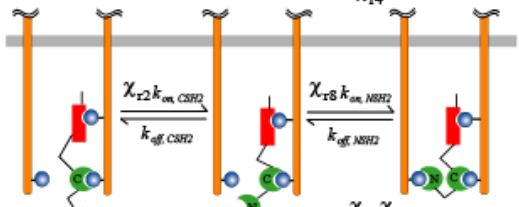




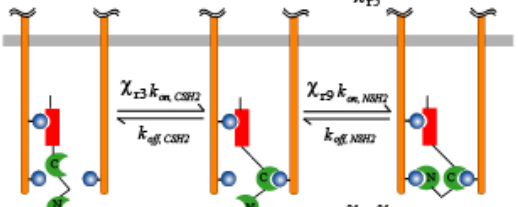
$$\chi_{r10} = \frac{\chi_{r1}\chi_{r6}}{\chi_{r4}}$$



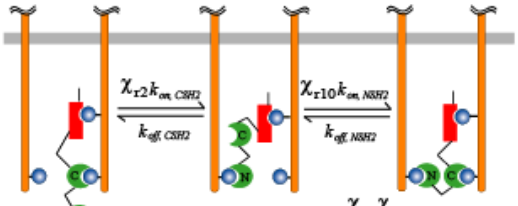
$$\chi_{r11} = \frac{\chi_{r1}\chi_{r7}}{\chi_{r5}}$$



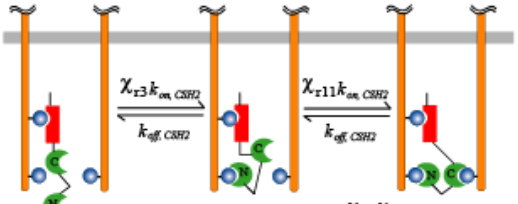
$$\chi_{r8} = \frac{\chi_{r1}\chi_{r6}}{\chi_{r2}}$$



$$\chi_{r9} = \frac{\chi_{r1}\chi_{r7}}{\chi_{r3}}$$



$$\chi_{r10} = \frac{\chi_{r1}\chi_{r6}}{\chi_{r4}}$$



$$\chi_{r11} = \frac{\chi_{r1}\chi_{r7}}{\chi_{r5}}$$

APPENDIX B

B1 Mathematical expression for effective dissociation constant in bivalent tandem interactions

(Refer to Eq. 4.1, Chapter 4).

Consider a protein molecule S having two SH2 domains namely N-SH2 and C-SH2, which bind with distinct affinities with two phosphorylated tyrosine residues (pY1 and pY2) of a peptide. The biding affinities of these two protein domains for pY1 site are $K_{D,N1}$ and $K_{D,C1}$, and for pY2 site are $K_{D,N2}$ and $K_{D,C2}$, respectively. Assume that this tandem interaction involves a cooperativity (arising from intracomplex localization effect of partner domains), and the enhancement factor χ quantifies the extent of this cooperativity.

The peptide – protein interactions can be written as follows. The enhancement/ cooperativity factor (χ) signifies a ring-closure interaction in each case.

$$R + S \underset{k_{r,N1}}{\overset{k_{f,N1}}{\longleftrightarrow}} C_{N1} \underset{k_{r,C2}}{\overset{\chi k_{f,C2}}{\longleftrightarrow}} C_{N1C2}$$

$$R + S \underset{k_{r,N2}}{\overset{k_{f,N2}}{\longleftrightarrow}} C_{N2} \underset{k_{r,C1}}{\overset{\chi k_{f,C1}}{\longleftrightarrow}} C_{N2C1}$$

$$R + S \underset{k_{r,C1}}{\overset{k_{f,C1}}{\longleftrightarrow}} C_{C1} \underset{k_{r,N2}}{\overset{\chi k_{f,N2}}{\longleftrightarrow}} C_{N2C1}$$

$$R + S \underset{k_{r,C2}}{\overset{k_{f,C2}}{\longleftrightarrow}} C_{C2} \underset{k_{r,N1}}{\overset{\chi k_{f,N1}}{\longleftrightarrow}} C_{N1C2}$$

Dissociation constants in terms of the association and dissociating rate constants are:

$$K_{D,N1} = \frac{k_{r,N1}}{k_{f,N1}};\ K_{D,N2} = \frac{k_{r,N2}}{k_{f,N2}};\ K_{D,C1} = \frac{k_{r,C1}}{k_{f,C1}};\ K_{D,C2} = \frac{k_{r,C2}}{k_{f,C2}}$$

Steady-state mass balance equations (B1.1 through B1.6)

$$\frac{d[C_{N1}]}{dt} = k_{f,N1}[R][S] + k_{r,C2}[C_{N1C2}] - \left(k_{r,N1} + \chi k_{f,C2}\right)[C_{N1}] = 0 \qquad \text{B1.1}$$

$$\frac{d[C_{N2}]}{dt} = k_{f,N2}[R][S] + k_{r,C1}[C_{N2C1}] - \left(k_{r,N2} + \chi k_{f,C1}\right)[C_{N2}] = 0 \qquad \text{B1.2}$$

$$\frac{d[C_{C1}]}{dt} = k_{f,C1}[R][S] + k_{r,N2}[C_{N2C1}] - \left(k_{r,C1} + \chi k_{f,N2}\right)[C_{C1}] = 0 \qquad \text{B1.3}$$

$$\frac{d[C_{C2}]}{dt} = k_{f,C2}[R][S] + k_{r,N1}[C_{N1C2}] - \left(k_{r,C2} + \chi k_{f,N1}\right)[C_{C2}] = 0 \qquad \text{B1.4}$$

$$\frac{d[C_{N1C2}]}{dt} = \chi\left(k_{f,C2}[C_{N1}] + k_{f,N1}[C_{C2}]\right) - \left(k_{r,C2} + k_{r,N1}\right)[C_{N1C2}] = 0 \qquad \text{B1.5}$$

$$\frac{d[C_{N2C1}]}{dt} = \chi\left(k_{f,C1}[C_{N2}] + k_{f,N2}[C_{C1}]\right) - \left(k_{r,C1} + k_{r,N2}\right)[C_{N2C1}] = 0 \qquad \text{B1.6}$$

$$\text{B1.1} \Rightarrow [C_{N1}] = \frac{k_{f,N1}[R][S] + k_{r,C2}[C_{N1C2}]}{k_{r,N1} + \chi k_{f,C2}} \qquad \text{B1.7}$$

$$\text{B1.2} \Rightarrow [C_{N2}] = \frac{k_{f,N2}[R][S] + k_{r,C1}[C_{N2C1}]}{k_{r,N2} + \chi k_{f,C1}} \qquad \text{B1.8}$$

$$\text{B1.3} \Rightarrow [C_{C1}] = \frac{k_{f,C1}[R][S] + k_{r,N2}[C_{N2C1}]}{k_{r,C1}[C_{C1}] + \chi k_{f,N2}[C_{C1}]} \qquad \text{B1.9}$$

$$\text{B1.4} \Rightarrow [C_{C2}] = \frac{k_{f,C2}[R][S] + k_{r,N1}[C_{N1C2}]}{k_{r,C2} + \chi k_{f,N1}} \qquad \text{B1.10}$$

From B1.7 and B1.10,

$$\chi\left(k_{f,C2}[C_{N1}] + k_{f,N1}[C_{C2}]\right) = \chi k_{f,C2}\left\{\frac{k_{f,N1}[R][S] + k_{r,C2}[C_{N1C2}]}{k_{r,N1} + \chi k_{f,C2}}\right\} + \chi k_{f,N1}\left\{\frac{k_{f,C2}[R][S] + k_{r,N1}[C_{N1C2}]}{k_{r,C2} + \chi k_{f,N1}}\right\} \qquad \text{B1.11}$$

$$\Rightarrow \left[k_{r,C2}\left(1-\frac{\chi k_{f,C2}}{k_{r,N1}+\chi k_{f,C2}}\right)+k_{r,N1}\left(1-\frac{\chi k_{f,N1}}{k_{r,C2}+\chi k_{f,N1}}\right)\right][C_{N1C2}]=$$

$$\chi k_{f,C2}k_{f,N1}\left(\frac{1}{k_{r,N1}+\chi k_{f,C2}}+\frac{1}{k_{r,C2}+\chi k_{f,N1}}\right)[R][S]$$

$$\Rightarrow k_{r,C2}k_{r,N1}\left(\frac{1}{k_{r,N1}+\chi k_{f,C2}}+\frac{1}{k_{r,C2}+\chi k_{f,N1}}\right)[C_{N1C2}]=$$

$$\chi k_{f,C2}k_{f,N1}\left(\frac{1}{k_{r,N1}+\chi k_{f,C2}}+\frac{1}{k_{r,C2}+\chi k_{f,N1}}\right)[R][S]$$

$$\Rightarrow \frac{\chi[R][S]}{K_{D,C2}K_{D,N1}}=[C_{N1C2}] \qquad \text{B1.12}$$

Similarly,

$$\frac{\chi[R][S]}{K_{D,C1}K_{D,N2}}=[C_{N2C1}] \qquad \text{B1.13}$$

B1.12 + B1.13 $\Rightarrow$

$$\chi[R][S]\left\{\frac{1}{K_{D,C2}K_{D,N1}}+\frac{1}{K_{D,C1}K_{D,N2}}\right\}=[C_{N1C2}]+[C_{N2C1}]$$

$$\Rightarrow \frac{[R][S]}{[C_{N1C2}]+[C_{N2C1}]}=\frac{1}{\chi}\left\{\frac{1}{K_{D,C2}K_{D,N1}}+\frac{1}{K_{D,C1}K_{D,N2}}\right\}^{-1}$$

$$\therefore K_{D,eff.}=\frac{1}{\chi}\left\{\frac{1}{K_{D,C2}K_{D,N1}}+\frac{1}{K_{D,C1}K_{D,N2}}\right\}^{-1} \qquad \text{B1.14}$$

B2 Mathematical expression for protein concentration in solution in an surface plasmon resonance (SPR) inhibition assay

(refer to Eq. 4.2, Chapter 4)

Assume the context of an SPR inhibition experiment, where protein a soluble protein (*S*) binds to a peptide C1 immobilized on a solid surface. Anaother peptide /inhibitor, Ci in solution competes with the immobilized peptide for protein binding. In a typical experiment, the amount of the immobilized peptide should be very small. In the following derivation, $C_{1,T}$ C_T and S_T stand for total concentrations of the immobilized peptide, inhibitory peptide, and protein, respectively.

Reactions in the system (upon inhibitory peptide addition):

$$S + C_1 \underset{k_{r,C1}}{\overset{k_{f,C1}}{\longleftrightarrow}} SC_1 \qquad K_{D,C1} = \frac{k_{r,C1}}{k_{f,C1}}$$

$$S + C_i \underset{k_{r,Ci}}{\overset{k_{f,Ci}}{\longleftrightarrow}} SC_i \qquad K_{DCi} = \frac{k_{r,Ci}}{k_{f,Ci}}$$

Since concentration of immobilized peptide is very small:

$$K_{D,C1} = \frac{S_{free}[C_1]}{[SC_1]}$$

$$K_{D,C1}[SC_1] = \left(C_{1,T} - [SC_1]\right)S_{free}$$

$$[SC_1] = \frac{C_{1,T}S_{free}}{K_{D,C1} + S_{free}}$$

$$\text{Bound fraciton} = \frac{S_{free}}{K_{D,C1} + S_{free}} \qquad \text{B2.1}$$

S_{free} is the amount of protein not in complex with immobilized or inhibitory peptide

'Bound fraction' is the fraction of immobilized peptide in complex with the protein.

Now, considering the inhibitory peptide binding:

$$K_{DCi} = \frac{S_{free}[C_i]}{[SC_i]}$$

$$K_{DCi} = \frac{S_{free}[C_T - (S_T - S_{free})]}{(S_T - S_{free})}$$

$$S_{free}{}^2 - (S_T - C_T - K_{DCi})S_{free} - K_{DCi}S_T = 0$$

$$b = (S_T - C_T - K_{DCi}); \quad c = K_{DCi}S_T$$

$$\therefore S_{free}{}^2 + bS_{free} - c = 0$$

$$S_{free} = \frac{b + \sqrt{b^2 + 4c}}{2} \qquad \text{B2.2}$$

B3 Simulated isothermal titration calorimetry (ITC) data at χ values 1 μM to 100 μM

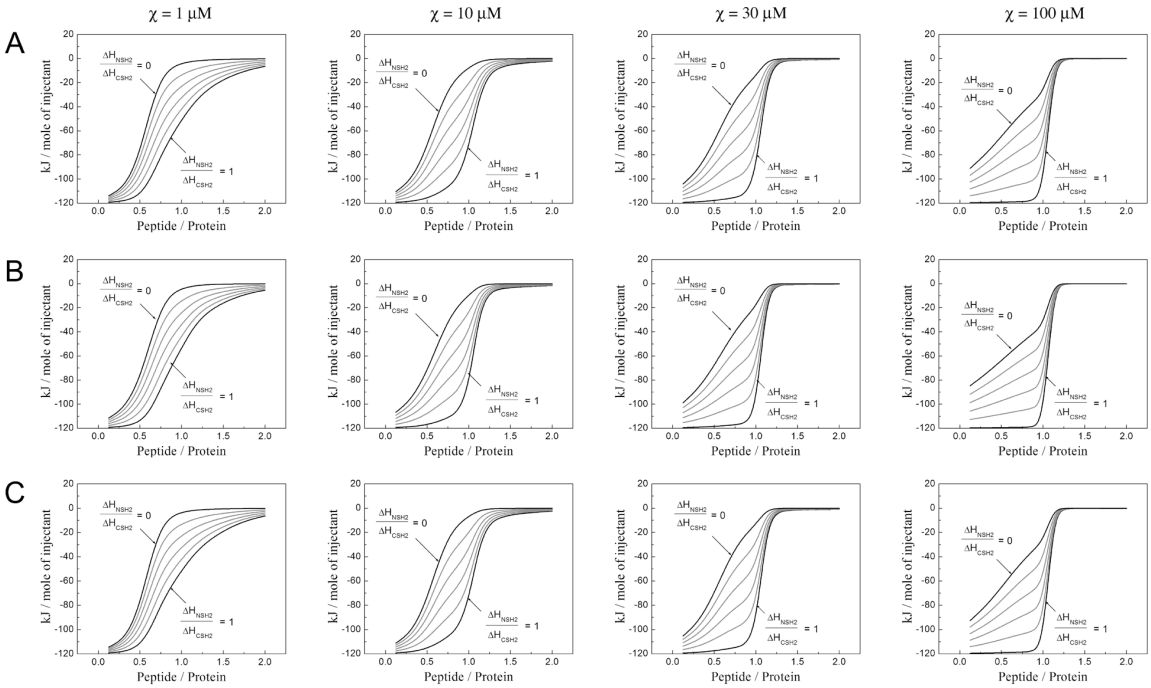


FIGURE B3 Further evaluation of isothermal titration calorimetry (ITC) experiments. Calculations were performed as described in the Fig. 4.4 caption, except that the values of χ used here better illustrate the transition between χ = 10 μM and χ = 100 μM. Hypothetical enthalpy changes with each injection of peptide were calculated, with ΔH_{CSH2} = –60 kJ/mol for C-SH2 bonds and various ratios of $\Delta H_{NSH2}/\Delta H_{CSH2}$; the curves in gray are with $\Delta H_{NSH2}/\Delta H_{CSH2}$ = 0.2, 0.4, 0.6, and 0.8. *A*. The same K_D values used in Fig. 4 ($K_{D,C1} = K_{D,C2}$ = 50 nM, $K_{D,N1} = K_{D,N2}$ = 1.5 μM) are reprised here. *B*. In this case, C-SH2 has a higher affinity for site 1, while N-SH2 has a slightly lower affinity for site 2 ($K_{D,C1}$ = 10 nM, $K_{D,C2}$ = 100 nM, $K_{D,N1}$ = 1 μM, $K_{D,N2}$ = 2 μM). *C*. Here again, C-SH2 has a higher affinity for site 1, and N-SH2 also has a slightly higher affinity site 2 ($K_{D,C1}$ = 10 nM, $K_{D,C2}$ = 100 nM, $K_{D,N1}$ = 2 μM, $K_{D,N2}$ = 1 μM).

B4 Concentration profiles of certain groups of complexes in a simulated isothermal titration calorimetry (ITC) experiments.

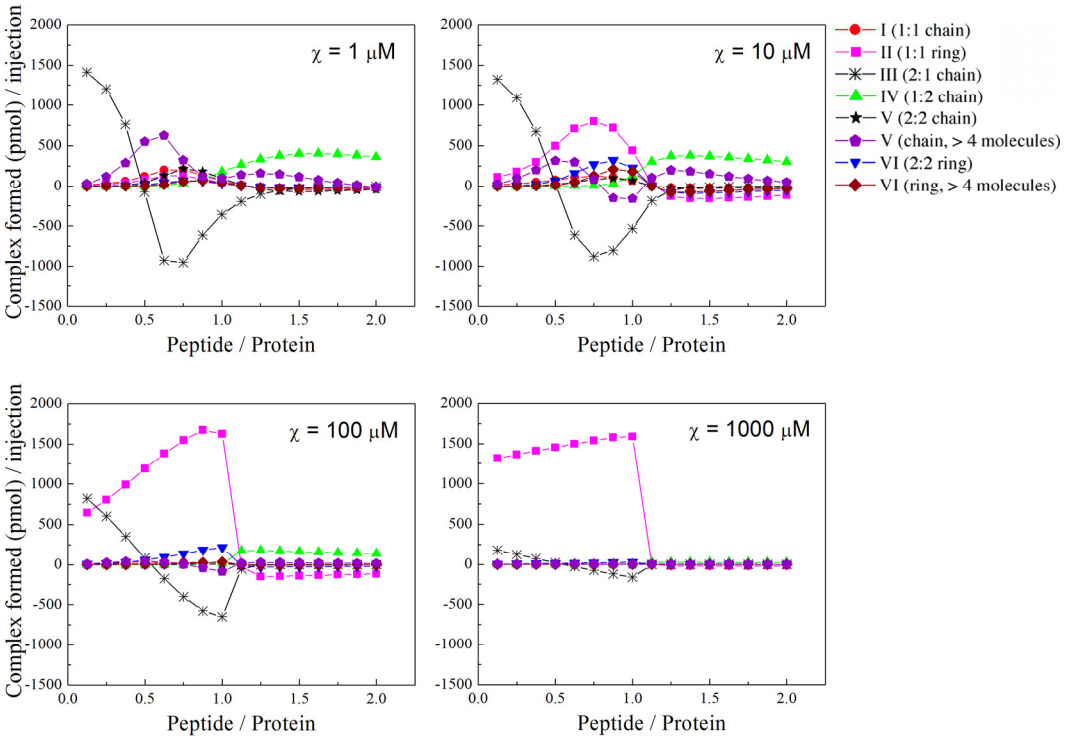


FIGURE B4 Analysis of complexes formed during a hypothetical isothermal titration calorimetry (ITC) experiment. See the caption under Fig. 4.4 for description of the experimental conditions. Structural classification of complexes is depicted in Fig. 4.1. The amount of each complex type is expressed in terms of the net change in abundance (in pmol) after each injection; negative values indicate a net reduction. As the cooperativity parameter, χ, increases, type II ring structures become increasingly dominant.