# Predicting Biological Networks from Diverse Data: a New Role for the Hsp90 Complex

**Chad L. Myers[1,2], Camelia Chiriac[1,2], Kara Dolinski[2], Olga G. Troyanskaya[1,2]**

[1]Department of Computer Science and [2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544.

Heat-shock proteins (HSP's) are a general class of proteins that function as molecular chaperones, ensuring proper folding of other proteins particularly under stressful conditions. Most heat shock proteins are constitutively expressed but over-expressed when the cell experiences heat shock or other stresses. One of the most abundant HSP's is the Hsp90 family, which is highly conserved from bacteria to humans and comprises about 1-2% of the total protein in a cell under normal conditions and 4-6% under cellular stress [1]. Hsp90 is a distinctive heat shock protein in that rather than functioning as a general cytosolic chaperone, most of its clients are signaling kinases or hormone receptors [2]. In humans, Hsp90 has recently drawn considerable attention, because it has shown promise as an anti-cancer target. Several compounds have been developed to inhibit the activity of Hsp90, which results in the rapid degradation of cell signaling and growth proteins, thus making it an effective tumor suppressor [1]. While such drugs have produced desirable effects in preliminary trials, little is known about the specific role Hsp90 plays in the function of each of these signaling kinases.

Hsp90's involvement in many essential biological processes has made it the focus of several recent high-throughput assays. The wealth of diverse data from those experiments undoubtedly holds important clues to Hsp90's function, yet harnessing this information presents a challenge for the traditional biologist. We have developed a general system for robust, process-specific integration of diverse genomic data and prediction of biological pathways. The basis of our approach is Bayesian learning for data integration followed by a probabilistic graph search that performs biological network prediction. We have measured considerable variation in the relevance and reliability of common experimental technologies across biological processes and have found that exploiting this variation by targeting integration and search toward the specific biological context of interest can significantly improve prediction accuracy. We have implemented this methodology in a public, web-accessible system for *Saccharomyces cerevisiae* (baker's yeast) and currently include genomic evidence from more than 6500 publications. Yeast is an attractive organism for such a system because experimental genomic data is readily available, and furthermore, yeast is an ideal model for studying eukaryotic processes.

We have used our context-sensitive integration and prediction system to analyze the Hsp90 homologs and associated co-chaperones in yeast. Surprisingly, we predict cross-talk between several Hsp90 proteins and members of a complex involved in DNA replication, Dbf4 and Cdc7. Specifically, the Cdc7-Dbf4 kinase phosphorylates the MCM complex at origins of replication, initiating the process of DNA replication [3]. Such a link between Hsp90 and proteins directly involved in regulating ploidy has not been previously characterized, and is particularly intriguing because abnormality in DNA content is known to play a central role in tumorigenesis and cancer progression.

We experimentally confirmed this predicted link to DNA replication by performing a number of genetic interaction experiments with Cdc7 or Dbf4 and members of the Hsp90 complex. These experiments revealed several previously unreported synthetic lethal interactions between Hsp90 genes and CDC7 and DBF4. We also found that several Hsp90 genes and co-chaperones suppressed a replication-related lethal phenotype, but in doing so, caused severe ploidy abnormalities, suggesting the mutants underwent re-replication. We have gained additional insight from microarray experiments characterizing the expression of Hsp90 and DNA replication genes for several of the related double and triple mutants. These results confirm the predicted link between Hsp90 and DNA replication, and experiments are ongoing to further characterize Hsp90's role in controlling DNA replication.

Thus, our computational predictions and experimental confirmation offer insight on the mechanism by which compounds targeting Hsp90 are effective in treating cancer. Furthermore, this example illustrates the power of computational tools tailored toward studying specific biological processes. An integrated view of the experimental data and a directed, context-sensitive approach to network prediction are an effective combination for constructing models that are specific and accurate enough to direct laboratory experiments.

[1] Goetz MP, Toft DO, Ames MM, Erlichman C (2003) The Hsp90 chaperone complex as a novel target for cancer therapy. *Ann Oncol.* 14:1169-76.
[2] Young JC, Moarefi I, Hartl FU (2001) Hsp90: a specialized but essential protein-folding tool. *J Cell Biol.* 154:267-73.
[3] Cheng L, Collyer T, Hardy CFJ (1999) Cell cycle regulation of DNA replication initiator factor Dbf4p. *Mol Cell Biol.* 19:4270-8.