

# RICHARD BONNEAU

## Assistant Professor

New York University, Center for Genomics and Systems Biology  
12 Waverley, New York, NY 10003, USA  
[bonneau@nyu.edu](mailto:bonneau@nyu.edu), 212-992-9516

### Education

#### **Ph.D. in Biochemistry, 2001**

Biomolecular Structure and Design  
University of Washington, Seattle, WA,

#### **BA in Biochemistry, magna cum Laude, 1997**

Florida State University, Tallahassee, FL

### Current Appointments

#### **Assistant Professor 2005-current**

New York University, Center for Genomics & Systems Biology  
Department of Biology (Primary appointment)  
Department of Computer Sciences (Courant Institute, Joint appointment)

#### **Affiliate Faculty 2006-current.**

NYU/NIH Graduate Partnership Program in Structural Biology  
(Skirball Institute / NYU Langone Medical Center).

#### **Affiliate Professor 2007-current.**

Institute for Systems Biology, Seattle.

### Previous Appointments

#### **Senior Advisor, Computational Biology Tacitus, LLC, Philadelphia, PA, USA 2007-2009.**

With Howard Coale, Using Video Game technology and software libraries to visualize biological data (so-called data-gaming).

#### **Sr. Scientist. 2002-2005. Institute for Systems Biology, Seattle.**

With Dr. Leroy Hood. Network Inference, development of combined biclustering and motif detection methods; genome wide *de novo* structure prediction and structure-based genome annotation.

#### **Sr. Scientist. 2002. Structural GenomiX, San Diego (acquired by Lilly).**

Modeling and functional annotation of human kinase non-catalytic domains. Annotation of human secreted proteins of unknown function. X-ray crystallographic refinement-methods development.

#### **HHMI Pre-doctoral fellow. 1997-2001. University of Washington.**

With Dr. David Baker. *Ab initio* protein structure prediction methods development (Rosetta).

#### **Research Assistant. 1997. National High Magnetic Field Laboratory, Tallahassee.**

With Dr. Allen G. Marshall. Design of ultra low flow rate chromatography systems for proteome analysis of low concentration biological samples via Fourier Transform Ion Cyclotron Resonance mass spectrometry.

#### **Undergraduate Research. Florida State University.**

1995-1996. With Dr. Tim Logan. Protein Folding, Study of FKBP-12 by proton exchange using NMR (MEXICO). Methods for isotopic labeling in mammalian cell lines.

1993-1994. With Dr George Bates. Plant transformation and genetic manipulation, plant cell culture.

### Honors and Awards

Discover Magazine: Top 20 Scientists under 40, 2008

HHMI Pre-doctoral Fellowship in the Biological Sciences, 1998-2001

Magna Cum Laude, Biochemistry, FSU, 1996

American Cancer Society – James Jay Fisher Fellowship, 1996  
Florida Academic Scholars Award, 1993

## Publications

1. Waltman P, Kacmarczyk T, Bate AR, Kearns DB, Reiss DJ, Eichenberger P, **Bonneau R**. Multi-species integrative biclustering. *Genome Biol.* 2010;11(9):R96. Epub 2010, Sep 29. PubMed PMID: 20920250

*Significance and role:* This paper described a new algorithm for biclustering multi-species genomics data sets. Our first application of this algorithm revealed a surprising number of conserved co-regulated modules across a dataset spanning the gram positive bacteria. We compared our algorithm to a large number of other biclustering and clustering algorithms and demonstrate improved performance. We show that this improved performance is due to the comparative (multiple species) and integrative (multiple data types) nature of our algorithm. I conceived of, implemented and extensively tested a new algorithm for finding conserved co-regulated groups in multiple species datasets.

2. Greenfield A, Madar, A, **Bonneau R**. DREAM4: Combining Genetic and Dynamic Information to Identify Biological Networks and Dynamical Models. DREAM4 top performers special collection. *PLoS ONE* 2010, 5(10): e13397. doi:10.1371/journal.pone.0013397

*Significance and role:* This is an invited paper in the DREAM4 special collection that describes our top-performing regulatory network inference method which my group developed (co-1st in the large-network inference challenge). This paper describes both our core method, new ways of estimating the error distributions for network parameters and new ways of combining genetic data (genomic analysis of knock outs and over expression) with time series data. Our use of knock-out and time-series information allowed us to predict both the network topology, the result of double knock-outs, and transcript levels at future time-points given initial conditions (our use of knock out information still results in ODE models useful in modeling system dynamics).

3. Miao EA, Mao DP, Yudkovsky N, **Bonneau R**, Lorang CG, Warren SE, Leaf IA, Aderem A. Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. *Proc Natl Acad Sci U S A.* 2010 Feb 16;107(7):3076-80. Epub 2010 Feb 1. PubMed PMID: 20133635.

*Significance and role:* In this collaborative work we provided structure predictions of newly discovered proteins participating in the type III secretion pathway (a pathway important for several bacterial pathogens). My lab used structure prediction to detect similarities between the structure of these proteins and proteins in making up the human and mouse innate immune systems (not apparent by sequence alone).

4. Madar A, Greenfield A, Vanden-Eijnden E, **Bonneau R**. DREAM3: Network Inference Using Dynamic Context Likelihood of Relatedness and the Inferelator. *PLoS ONE*, 2010, 5(3): e9803. doi:10.1371/journal.pone.0009803

*Significance and role:* This is an invited paper in the DREAM3 top-performers PLoS special collection. In this paper we focused on optimizing our performance without explicitly using knock-out information. We show that a small modification to an information theory based method (called dynamic-CLR) that allows for explicit use of time series data greatly improves the previously published CLR method (dynamic-CLR outperforms CLR) and our results when combining this new method with the Inferelator. We examine new, more scalable approaches to model selection in the context of large network inference.

5. Madar A, Greenfield A, Ostrer H, Vanden-Eijnden E, **Bonneau R**. The Inferelator 2.0: A scalable framework for reconstruction of dynamic regulatory network models. *Conf Proc IEEE Eng Med Biol Soc.* 2009;1:5448-51. PubMed PMID: 19964678.

*Significance and role:* In this work my lab developed new methods for refining large differential equation models of transcriptional regulatory network dynamics and topology. We present, implement, and test a new Markov-Chain- Monte-Carlo (MCMC) network refinement method, Inferelator 2.0, that works in tandem with Inferelator 1.0. We show results for the prokaryote *Halobacterium* that demonstrate a marked improvement in our predictive performance in modeling the regulatory dynamics of the system

over longer time-scales.

6. Butterfoss GL, Renfrew PD, Kuhlman B, Kirshenbaum K, **Bonneau R**. A preliminary survey of the peptoid folding landscape. *J Am Chem Soc*. 2009 Nov 25;131(46):16798-807. PubMed PMID: 19919145.

*Significance and role:* We present an analysis of the conformational preferences of N-substituted glycine peptoid oligomers (a popular chemistry for peptide like oligomers where the side chain branches from the backbone nitrogen instead of the alpha carbon). We survey the backbone conformations observed in experimentally determined peptoid structures and provide a comparison with high-level quantum mechanics calculations of short peptoid oligomers. This paper lays the groundwork for incorporating peptoid side chains and peptoid backbones into Rosetta-design, which will in turn enable rational design of non-natural peptoid scaffolds and foldamers.

7. Wang KH, Isidro AL, Domingues L, Eskandarian HA, McKenney PT, Drew K, Grabowski P, Chua MH, Barry SN, Guan M, **Bonneau R**, Henriques AO, Eichenberger P. The coat morphogenetic protein SpoVID is necessary for spore encasement in *Bacillus subtilis*. *Mol Microbiol*. 2009 Nov;74(3):634-49. Epub 2009 Sep 22. PubMed PMID: 19775244; PubMed Central PMCID: PMC2806667.

*Significance and role:* My group used structure prediction (fold recognition and Rosetta de novo structure prediction) to suggest functions and evolutionary origin for key proteins involved in spore formation in *B. subtilis*.

8. Konieczka JH, Drew K, Pine A, Belasco K, Davey S, Yatskievych TA, **Bonneau R**, Antin PB. BioNetBuilder2.0: bringing systems biology to chicken and other model organisms. *BMC Genomics*. 2009 Jul 14;10 Suppl 2:S6. PubMed PMID: 19607657; PubMed Central PMCID: PMC2709267.

*Significance and role:* This paper describes an update to our BionetBuilder code, an open source code developed in my laboratory. Bionetbuilder has many users (5,000 downloads from our site; the plugin is also on the Cytoscape web page). Many of those users want to work in less well studied organisms and newly sequenced organisms. This extension adds the ability to use orthology relationships to map interactions to these less well studied organisms (not novel but very useful). My lab helped the student from the Antin lab who applied the code to a study of Chicken genomics data understand the BionetBuilder code, provided BionetBuilder 1.0 source code, and helped write the paper.

9. Whitehead K, Pan M, Masumura K, **Bonneau R**, Baliga NS. Diurnally entrained anticipatory behavior in archaea. *PLoS One*. 2009;4(5):e5485. Epub 2009 May 8. PubMed PMID: 19424498; PubMed Central PMCID: PMC2675056.

*Significance and role:* This paper describes the detection of a circadian rhythm in an Archea, a first. This work was primarily done in a collaborating lab, I performed Lomb-Scargle periodogram analysis, and other methods to detect and estimate the significance of cyclic expression.

10. Shah NH, Butterfoss GL, Nguyen K, Yoo B, **Bonneau R**, Rabenstein DL, Kirshenbaum K. Oligo(N-aryl glycines): a new twist on structured peptoids. *J Am Chem Soc*. 2008 Dec 10;130(49):16622-32. PubMed PMID: 19049458.

*Significance and role:* This work is very related to the work in publication 6, above, and represents the beginning of what is now a tight collaboration between my lab and the laboratory of Kent Kirshenbaum at NYU chemistry. My lab performed QM and MM calculations to look at how side chain choice (in this case N-aryl side chains and derivatives) effect back-bone degrees of freedom. This is a first step towards making peptoid backbone-dependent rotamer libraries that will be required if we are to incorporate peptoids into Rosetta-design.

11. Jochen Supper, Claas aufm Kampe, Dierk Wanke, Kenneth W. Berendzen, Klaus Harter, **Richard Bonneau**, and Andreas Zell. Modeling gene regulation and spatial organization of sequence based motifs. 8th IEEE international conference on BioInformatics and BioEngineering (BIBE 2008).

*Significance and role:* Jochen (now Dr. Supper) was an interdoc in my lab visiting from the lab of Dr. Andreas Zell. My lab helped to conceive part of this project and Jochen implemented the first part of this work while in my lab, finishing the work upon his return to Germany. This work examines how spatial organization of binding site motifs can be incorporated into the characterization of transcriptional

modules, and showed that accounting for spatial organization of binding sites is key to determining which sites are functional.

12. Mike Boxem, Zoltan Maliga, Niels J. Klitgord, Na Li, Irma Lemmens, Miyeko Mana, Lorenzo De Lichtervelde, Joram Mul, Diederik van de Peut, Maxime Devos, Nicolas Si-monis, Anne-Lore Schlaitz, Murat Cokol, Muhammed A. Yildirim, Tong Hao, Changyu Fan, Chenwei Lin, Mike Tipsword, Kevin Drew, Matilde Galli, Kahn Rhrissorakrai, David Drechsel, David E. Hill, **Richard Bonneau**, Kristin C. Gunsalus, Frederick P. Roth, Fabio Piano, Jan Tavernier, Sander van den Heuvel, Anthony A. Hyman, Marc Vidal. A highly connected domain based interaction network of *C. elegans* early embryogenesis. **Cell**, 2008 134(3) pp. 534 - 545.

*Significance and role:* This work assayed interactions between domains excised from whole proteins in *C. elegans*. Each protein in this large genomic experiment was first cut into several domains. My group then explored the degree to which the interactions observed mapped onto structurally characterized domains and interactions, finding a good correspondence between structural domains and minimal-interaction domains (in many cases we could make structural models of either the individual domains or the interaction, in other cases we relied on simpler methods for predicting protein domain boundaries).

13. **Bonneau, R\***, Facciotti, MT, Reiss, DJ, Madar A, et al., Baliga, NS\*. A predictive model for transcriptional control of physiology in a free living cell. (2007) **Cell**. Dec 131:1354-1365

*Significance and role:* This paper described a large-scale genomics effort and several computational advances that led to a differential equation model of the transcriptional regulatory dynamics of over 80% of the *Halobacterium* genome. My group developed the algorithms and modeling techniques that made the main result of this paper; I also participated in several aspects of the experimental design, helped conceive of and implement codes used to visualize the data, was one of the two authors that wrote the paper, and have since continued to develop and distribute the codes I developed as part of this work. [\* *co-first and co-corresponding author*]

14. Iliana Avila-Campillo\*, Kevin Drew\*, John Lin, David J. Reiss, **Richard Bonneau**. BioNetBuilder, an automatic network interface. (2007) **Bioinformatics**. Feb 1;23(3):392-3.

*Significance and role:* There are several tools for genomic data analysis and often the barrier to use is getting data into the tools. This paper describes a simple plugin for the network visualization tool Cytoscape that we developed that gives biologists several ways of getting network data into Cytoscape. The tool also interfaces with the Gagggle, a tool for connecting this tool to a large number of other Gagggle-enabled tools and websites.

15. Andersen-Nissen E, Smith KD, **Bonneau R**, Strong RK, Aderem A. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. (2007) **J Exp Med**. Feb 19;204(2):393-403.

*Significance and role:* This work was primarily carried out in the lab of Alan Aderem, our lab provided bioinformatic support. Our main contribution was to provide a structure prediction that was of sufficient quality to redesign the mouse TLR5 to better model a human TLR5 binding to flagellin from various bacterial species. The structure was used to help identify the binding site and then reengineer the binding site to resemble the human TLR5 binding site. The recent crystal structure for portion of TLR5 shows that our prediction was of very high quality in the 200 residues surrounding the binding site).

16. Malmström L, Riffle M., Strauss CEM, Chivian, D, Davis TN., **Bonneau R.**<sup>3</sup> Baker D.: Genome-wide superfamily assignments for *Saccharomyces cerevisiae* protein domains through integration of *de novo* structure prediction with the gene ontology. **PLoS Biol.** (2007) Apr;5(4):e76.

*Significance and role:* This work uses structure prediction (Rosetta *de novo* and fold recognition) to predict the fold for > 1000 protein domains with no prior structure annotation in the yeast genome. We show that integration of predicted fold with functional information contained in the Gene Ontology database greatly improves accuracy, coverage and interpretability of our results. My lab contributed to all aspects of this work, from Rosetta core development to carrying out the predictions on the World Community Grid (via a collaboration between my lab and IBM).

17. Facciotti MT, Reiss DJ, Pan M, Kaur A, Vuthoori M, **Bonneau R**, Shannon P, Srivastava A,

Donohoe SM, Hood LE, Baliga NS. General transcription factor specified global gene regulation in archaea. **Proc Natl Acad Sci U S A**. (2007) Mar 13;104(11):4630-5.

*Significance and role:* This work used ChIP-chip to examine a large family of paralogous general transcription factors in Halobacterium and found that they act as master regulators (in a manner similar to sigma factors in bacteria). I aided in analysis of ChIP-chip data and also provided predicted regulatory network and co-regulated clusters used to interpret the data.

18. Kye MJ, Liu T, Levy SF, Xu NL, Groves BB, **Bonneau R**, Lao K, Kosik KS. Somatodendritic microRNAs identified by laser capture and multiplex RT-PCR. **RNA**. (2007) Aug;13(8):1224-34.

*Significance and role:* This work found evidence of active transport in neurons of microRNAs by assaying for ~100 microRNAs in neurite tips, cell bodies, and glial cells using laser dissection. My lab helped with a simple analysis to estimate the significance of the differences in microRNA concentration in samples taken from different regions of different neurons and glial cells.

19. **Bonneau R**, Reiss DJ, Shannon P, Hood L, Baliga NS, Thorsson V (2006) The Inferelator: a procedure for learning parsimonious regulatory networks from systems-biology data-sets de novo. **Genome Biol**. 7(5):R36.

*Significance and role:* This paper describes an algorithm I developed with Vestinn Thorsson called the Inferelator that learns simplified ODE models of transcriptional regulation from heterogeneous data collections (data collections composed of time series, knock-outs, environmental perturbations, etc.). This work is highly cited and led to the work that resulted in publication 13 (Bonneau et. al, Cell, 2007) above. This paper showed that explicit treatment of time could improve the accuracy and utility of learned networks even in the context of high-noise biological data-sets.

20. David J Reiss, Nitin S Baliga, **Bonneau R**. (2006) Integrated biclustering of heterogeneous genome-wide datasets. **BMC Bioinformatics**. 7(1):280.

*Significance and role:* This work describes an integrative biclustering algorithm cMonkey, co-developed with David Reiss. The paper shows that integration of network, binding site motifs, and expression data provides better estimation of condition dependent (environment dependent) co-regulate modules. We compared our algorithm to several other algorithms and showed that our integration of additional data-types was essential to performance. We also showed our algorithm to be superior for pre-clustering prior to network inference (further suggesting that our detected co-regulated groups are more accurate than those attainable with other tools available at the time).

21. Shannon P, Reiss DJ, **Bonneau R**, Baliga NS (2006) The Gaggle: A system for integrating bioinformatics and computational biology software and data sources. **BMC Bioinformatics**. 7:176.

*Significance and role:* This paper describes a tool for connecting different data visualization tools via a simple API (simple enough that ~50 tools and websites have been Gaggle-enabled to date). I participated in the design and testing of the tool.

#### **Prior to NYU:**

22. Flory MR, Lee H, **Bonneau R**, Mallick P, Serikawa K, Goodlett D, Morris D, Aebersold R. (2006) Quantitative proteomic analysis of the budding yeast cell cycle using acid-cleavable isotope-coded affinity tag reagents. **Proteomics** Dec;6(23):6146-57

*Significance and role:* In this paper microarray and proteomic experiments are collected at several time-points along the yeast cell cycle. I aided in the simultaneous analysis of proteomic and microarray data.

23. Zhang, H., Loriaux, P., Eng, J., **Bonneau R**, Smith, R & Aebersold, R. UniPep, a database for human N-linked glycosites: A resource for biomarker discovery. **Genome Biology**, (2006), 7:114

*Significance and role:* This paper describes new methods for large scale proteomic analysis of glycoproteins. I aided in the bioinformatic analysis required for this paper, including but not limited to prediction of trans-membrane regions, disordered regions, signal sequences and protein domain architecture. This allowed separation of observed glycoproteins into classes based on expected transport (membrane, secreted, etc.) and into distinct protein domains.

24. **Bonneau R**, Baliga NS, Deutsch EW, Shannon P, Hood L. (2004) Comprehensive de novo structure prediction in a systems-biology context for the archaea Halobacterium sp. NRC-1. **Genome Biology**. 5(8):R52-68  
*Significance and role:* In this work we used Ginzu (a tool I co-developed in the Baker lab) protein domain prediction to identify protein domains in Halobacterium. Domains without function were then predicted using Rosetta and fold recognition. We found reliable fold predictions for an additional 19% of the protein domains predicted for this organism and explored three sets of predictions (new transcription factors, newly annotated prophage regions, and new annotations for the chemotaxis and sensory systems for this cell). This was the first application of Rosetta *de novo* structure prediction to a whole proteome.
25. Baliga NS, Bjork SJ, **Bonneau R**, Pan M, Iloanusi C, Kottmann MC, Hood L, DiRuggiero J. (2004) Systems level insights into the stress response to UV radiation in the halophilic archaeon Halobacterium NRC-1. **Genome Res**. 14(6):1025-35. Epub 2004 May 12.  
*Significance and role:* For this paper I made an early version of the Halobacterium association network and aided in the clustering and interpretation of the microarray data. This paper explores the transcriptome wide response of Halobacterium to UV stress.
26. Baliga NS., **Bonneau R.**, Facciotti M, Pan M, Deutsch E, Glusman G, Shannon P, Chiu Y, Weng RS, Kan JR, Hung P, Date S, Marcotte E, Hood L, Ng V. (2004) Genome sequence of Haloarcula marismortui—a halophilic archaeon from the Dead Sea. **Genome Res**. 5(8):R52  
*Significance and role:* This paper describes the sequencing and annotation of a new Halophilic Archaeon genome. I annotated the genome (using sequence and structure-based tools) and helped build a network to interpret these annotations (operon predictions, co-evolutionary edges and predicted proximity in metabolic pathways).
29. Chivian D, Kim DE, Malmstrom L, Bradley P, Robertson T, Murphy P, Strauss CE, **Bonneau R**, Rohl CA, Baker D. (2003) Automated prediction of CASP-5 structures using the Robetta server. **Proteins**. 53 Suppl 6:524-33.  
*Significance and role:* This paper is an invited paper to describe our top performing structure and domain prediction server, Robetta. This paper describes a server used to participate in the automated portion of CASP5 (servers only, no manual tweaking). The protocol used by the server closely follows the protocol we used in CASP4 and was successful in both domain and structure prediction.
30. Tsai J, **Bonneau R**, Morozov AV, Kuhlman B, Rohl CA, Baker D. (2003) An improved protein decoy set for testing energy functions for protein structure prediction. **Proteins**. 53(1):76-87.  
*Significance and role:* Testing protein structure prediction requires generation of compact, but incorrect, structures. There are many ways to do this, this paper describes the generation of one such “decoy” set using Rosetta. Many of the structures in these decoy sets were generated by the protocols used in publication 28 and 27 above.
31. **Bonneau, R.**, Dylan Chivian, Charlie EM Strauss, Carol Rohl, David Baker. (2002) De Novo Prediction of Three Dimensional Structures for Major Protein Families. **JMB**, 322(1):65-78.  
*Significance and role:* This paper describes the generation of Rosetta structure predictions for large numbers of Pfam families that, at the time, lacked annotation and structure. Many of the predictions made in this paper have been validated since. This paper is also the first paper to describe a reliable confidence metric for Rosetta *de novo* structure prediction.
27. **Bonneau, R.**, Ingo Ruczinski, Jerry Tsai, David Baker. (2002) Contact Order and ab initio Structure Prediction. **Protein Science**, 11(8):1937-44.  
*Significance and role:* It has been observed that contact order is a surprisingly good predictor of protein folding rate (Contact order is the average sequence separation of all pairs of contacting residues in a protein structure; the higher the contact order the more complex the fold). In this paper I showed that contact order is also a very good predictor of the performance of Rosetta structure prediction (for

example, contact order is a much better predictor of Rosetta-performance than length). This paper explores this, and other, parallels between protein folding in the lab and *in silico*.

28. **Bonneau, R.**, Jerry Tsai, Ingo Ruczinski, Dylan Chivian, Carol Rohl, Charlie EM Strauss, David Baker. (2001) Rosetta in CASP4: Progress in *ab initio* protein structure prediction. **Proteins** 45(S5)119-126.

*Significance and role:* This paper is an invited paper in the CASP4 special issue and describes the performance of Rosetta in the “new folds” category (where “new folds” designates the prediction of protein folds without homology to experimentally determined protein structures). CASP is the Critical Assessment of Structure Prediction, a field wide test of methods, where predictors are given sequences that lack structure, and structures are revealed only after predictions are submitted to a secure server. We were able to correctly predict the folds for proteins in 16 of 22 proteins. In many cases we predicted the fold for multiple domains of multi-domain proteins correctly. This was the first use to domain prediction to split query sequences prior to analysis with Rosetta *de novo*. The protocol I developed to participate in CASP4 used parts developed from all the authors on the author list (I feel extremely lucky to have been part of that time). This protocol was later automated and turned into the Robetta server (publication 29).

32. Ruczinski, I., Kooperberg, C., **Bonneau, R.**, Baker, D. (2001) Distributions of beta sheets in proteins with application to structure prediction. **Proteins** 48(1), 85-97.

*Significance and role:* This paper is related to paper 27, above. I aided in the analysis by comparing my Contact Order results to results that Ingo Ruczinski obtained by comparing strand arrangements in Rosetta decoys to those found in native structures. I also generated structure predictions used in this paper (the large ensemble of Rosetta decoy conformations for several proteins).

33. **Bonneau, R.**, Jerry Tsai, Ingo Ruczinski, David Baker. (2001) Functional inferences from blind *ab initio* protein structure predictions. **J. Struct. Biol.** 134(2-3),186-90.

*Significance and role:* This paper is related to our CASP4 paper and shows that Rosetta *de novo* structures generated as part of CASP4 were in several cases sufficient to predict aspects of each protein’s function. This paper demonstrates the first function predictions based on *de novo* structure prediction (the term *ab initio* is used in the title but has been replaced by consensus with the term *de novo* to indicate that we do not use a template to make our predictions).

34. **Bonneau, R.**, Strauss, C. & Baker, D. (2001). Improving the performance of Rosetta using multiple sequence alignment information and global measures of hydrophobic core formation. **Proteins** 43(1), 1-11.

*Significance and role:* In this paper we show that combining Rosetta predictions for several members of a protein domain family can increase accuracy and the number of protein domains we can predict with Rosetta (yield). This paper remains relevant as sequencing efforts have uncovered several large (with hundreds of members) domain families of unknown function. This paper also describes a new score designed to reward contiguous well packed hydrophobic cores and penalize packings with high eccentricity. This paper is also the first description of the multiple-homolog nearest-neighbors clustering method that was the core of how we selected structure predictions for CASP4, and future CASPs.

35. Simons, K. T., **Bonneau, R.**, Ruczinski, I. & Baker, D. (1999). *Ab initio* protein structure prediction of CASP III targets using ROSETTA. **Proteins** 37(S3), 171-176.

*Significance and role:* This paper describes the first time Rosetta worked in a blind-prediction context. I worked with Kim Simons to make the predictions, developed parts of Rosetta, helped manually assess predictions, helped build the compute cluster that was used to make predictions, and developed a clustering algorithm used to select several of the top predictions we made that year (this paper demonstrated that scoring and selecting the best structures from large Rosetta ensembles is a critical problem).

**Book chapters, reviews and perspectives:**

36. **Bonneau, Richard**. Learning biological networks: from modules to dynamics. *Nature Chemical Biology* 4, 2008, 658 - 664
37. Waltman P, Kacmarczyk T, **Bonneau R**: Prokaryotic Systems Biology. In Plant Systems Biology, Annual Plant Reviews. 1st edition. Edited by: Coruzzi G, Gutierrez RA. Ames, Iowa: Blackwell Publishers; 2009:67-136
38. Madar A, **Bonneau R**. Learning global models of transcriptional regulatory networks from data. *Methods Mol Biol.* 2009;541:181. Review. PubMed PMID:19381524.
39. Kevin Drew, Dylan Chivian, **Richard Bonneau**. De novo structure prediction: methods and applications. *Structural Bioinformatics*, Second Edition. (2009), Edited by: Jenny Gu and Philip E. Bourne, p. 755- 781, John Wiley and Sons. **ISBN-10**: 0470181052
40. **Richard Bonneau**. (2008) Dissecting the Quorum-Sensing Receptor LuxN, *Cell*, 134(3):390-391, ISSN 0092-8674
41. Weston AD, Baliga NS, **Bonneau R**, Hood L. Systems approaches applied to the study of *Saccharomyces cerevisiae* and *Halobacterium* sp. *Cold Spring Harb Symp Quant Biol.* 2003;68:345-57. PubMed PMID: 15338636.
42. Chivian D, Robertson T, **Bonneau R**, Baker D. (2003) Ab initio methods. **Methods Biochem Anal.** 44:547-57.
43. **Bonneau, R** & Baker, D. (2001). Ab Initio Protein Structure Prediction: Progress and Prospects. **Annu. Rev. Biophys. Biomol. Struct.** 30, 173-89. Review.

#### **Manuscripts submitted / in preparation:**

44. Kevin Drew, Patrick Winters, Glenn L. Butterfoss, Viktors Berstis, Keith Uplinger, Jonathan Armstrong, Michael Riffle, Erik Schweighofer, Bill Bovermann, David R. Goodlett, Trisha N. Davis, Dennis Shasha<sup>6</sup>, Lars Malmström, **Richard Bonneau**. The proteome folding project: proteome-scale prediction of structure and function [ *submitted to Genome Research* ]
45. Christopher S. Poultney, David Gresham, Nathan J. Brandt, Glenn L. Butterfoss, Michelle R. Gutwein, Kevin Drew, Kristin C. Gunsalus, Dennis E. Shasha, **Richard Bonneau**. (2011) Structure based prediction of temperature sensitive mutations. [ *Invited submission to PLoS one special collection: RosettaCon 2010: Worked examples of macromolecular modeling and design* ]
46. Maria Ciofani, Aviv Madar, Fransis Kirigin, Rick Myeres\*, **Richard Bonneau\***, Dan Littman\*. A global transcription factor network underlying Th17 cell lineage specification. [In preparation]
47. Vincent A. Voelz, Glenn L. Butterfoss, Ilya Chorny, Ken A. Dill, Ron Zuckermann, Kent Kirshenbaum, Richard Bonneau.. Blind structure prediction of peptoid oligomers [In preparation]
48. Bishwajit Paul, Mikki G. Boswell, Fanny G. Yeung, Glenn L. Butterfoss, P. Douglas Renfrew, Christian Wolf, Richard Bonneau, Kent Kirshenbaum. [In preparation]
49. Peter Jordan, Glenn L. Butterfoss, P. Douglas Renfrew, Rich Bonneau, Kent Kirshenbaum. Oligo(N-alkoxy glycines): Trans substantiating peptoid conformation. [Submitted , Biopolymers]
50. Charlie EM Strauss, P. Douglas Renfrew, Richard Bonneau. The Rosetta Developers Meeting, 2010 special collection: state of the art macromolecular modeling meets reproducible publishing. [Invited, Invited introduction to the RosettaCon 2010 PLoS one special collection]
51. P. Douglas Renfrew, Tim Craven, Glenn L. Butterfoss, Richard Bonneau. Constructing rotamer libraries for the computational design of peptoid loldamers [In preparation]



52. Thadeous Kacmarczyk, Peter Waltman, Ashley R. Bate, Patrick Eichenberger, Richard Bonneau. The Comparative Microbial Modules Resource: Generation and Visualization of Multi-species Biclusters. [in preparation]

53. Melissa M. Pentony, Patrick Winter, Kevin Drew, Apurva Narechania, Sergios-Orestis Kolokotronis, Rob DeSalle, Richard Bonneau, Michael D. Purugganan. Integrating structural and sequence information to infer evolutionary history of plant gene families [In preparation]

## **Research Support**

Project/Proposal Title: "2010: The Arabidopsis folding project"

Source of Support: NSF

Total Award Amount: 1,621,765

Role on the Project: PI

CoPIs: Mike Purugganan (NYU) and Rob DeSalle (AMNH)

Project Period: 9/01/09 – 9/01/12 (we will submit a renewal of this grant in 2011)

Project title: Elucidation of the transcriptional network underlying the Th17 lineage program

Principal Investigators: Dan Littman, Richard Bonneau

Funding agency and grant number: NIH 1 RC4 AI092765-01

Role on project: PI

Total Award Amount: \$5,159,139

this grant provides \$156,272 / y to Dr. Bonneau for years 2010-2012

Project title: Role of ROR $\gamma$ t in the transcriptional regulatory network underlying Th17 lineage specification and function

PI: Dan Littman

Funding agency and grant Number: NIH RC1 AI087266-01

Role: Co-investigator

Dates: Recovery act funding 1y duration, 2010

This grant provides \$42,470 to Dr. Bonneau in 2010.

NIH Physical Sciences Oncology Center

Prime Institution: University of Southern California

Award/Contract#: 1 U54CA143907-01

Role: Co-investigator

Award amount: \$129,663 / y to Bonneau lab for five years (2010 = first year)

Period: 9/28/2009-7/31/2010

\*Total project period: 9/28/2009-7/31/2014

Institute for Systems Biology / Luxembourg Systems Biology Initiative:

Prime Institution: Institute for Systems Biology

Award/Contract#: 2009.0006-PO27428

Award amount: \$140,328 / y to Bonneau lab for five years (2010 = first year)

Project period: 07/01/2009-6/30/2010

NYU SoM Subcontract: NIH Nano Medicine Center (Mike Dustin, PI)

Award/Contract#: 7 PN2 EY016586-06

Award amount: \$200,000 / y to Bonneau lab for five years (until 2015)

Role: Co-investigator

Project period: 09/30/2009-9/29/2010 (the project was renewed for 5 years, until 2015)

### **Pending:**

NSF, Environmental Gene Regulatory Networks in Rice

Collaborative grant with Michael Purugganan

Role: Co-PI

Award Amount: 550,000 / y for 4 years (estimated, pending)  
Project Period: 9/1/2011 - 8/31/2015

## **Media-coverage of my work.**

My work was featured in the Dec. 2008 Issue of **Discover Magazine** as one of the top 20 scientists under 40. My work (the 2007 Cell paper on biological Network inference, on which I was 1st and co-corresponding author) was featured in over 100 articles, including Wired Magazine, Forbes, and several daily newspapers. My work on Protein Folding has also been widely covered in the media, including the Washington Post, New York Times, and Wired. The NYU torch is featured on over 200,000 screen-savers globally (on the desktops and laptops of people giving their unused CPU cycles to the human proteome folding project).

## **Speaking engagements.**

### **2011:**

January, 24th, Invited Seminar, Duke  
January, 9-10, Supercomputing: The Imperative and the Path Forward January 9-10, 2011 |  
InterContinental Hotel, Abu Dhabi

### **2010:**

December 9-10, 6th Annual meeting of the Nanomedicine Center for Mechanobiology, New York  
November 8, New York Academy of Sciences, Systems Biology of Bugs (one of three invited talks)  
September 7-10, Quantitative Biology: From Complex networks to simple models. Montauk, NY  
June, 18th, Physical Sciences Oncology, NIH  
May 12th, iCubed: Inflammation, Infection and Immunity retreat, New York  
March 28-30, EPSRC Symposium Workshop on Learning and inference in computational systems biology (LICSB), Warwick, England  
March 11, University of British Columbia, Invited Seminar  
February 25, UCSD Frontiers in Bioinformatics, Invited Seminar, (invited by UCSD graduate students)  
January 22-29. 9th Annual McGill Workshop on Bioinformatics, Barbados  
January 15, Invited Seminar, Cold Spring Harbor Laboratories  
January 5-8, Genomics and Systems Biology Workshop, Abu Dhabi

### **2009:**

December 16, Invited Seminar, Monsanto, St. Louis  
December 2-4, RECOMB-DREAM4-SB-RG, Invited top performer talk.  
November 20-22, National Academies, Keck-Futures Initiative workshop on synthetic biology.  
[resulted in publication of summaries from workshop: "NAKFI Synthetic Biology: Building a Nation's Inspiration" ]  
October 8-9, Berlin Summer meeting 2009, MDC Berlin  
September 21-24, EMBO: Practical Course on Networks in Biology analysis, modeling and reverse engineering , Bologna, Italy  
September 2-6, IEEE Engineering in Medicine and Biology Society (EMBC'09), Minneapolis  
August 5-7, RosettaCon 2010, the Rosetta Developers meeting, Speaker and co-organizer.  
August 9-10, 7th Peptoid Summit, [talk given by post-doctoral fellow, Glenn Butterfoss]  
July 10, IBM Watson Research Center, Invited Talk  
June 22-25, Statistical Modeling and Machine Learning in Computational Systems Biology, Tampere, Finland  
May 1-3, ZiF Workshop on Biological Networks, Bielefeld, Germany

## **2008:**

December 1st. Invited Talk. UCLA Bioinformatics Seminar Series.

November 3rd-4th. Keynote Speech. Robert-Cedergren Colloque (the Universite de Montreal Bioinformatics Meeting)

Oct 29-Nov 2, 2008. The 5th Annual RECOMB Satellite on Regulatory Genomics, the 4th Annual RECOMB Satellite on Systems Biology, and the 3rd Annual DREAM reverse engineering challenges, MIT.

Oct 22. MIT, The Microbial Systems Seminar at Parsons.

Oct 14-16. Invited Guest. NYU Medical/Skirball Annual Retreat.

September 4th. Invited Talk. Gene Network Sciences, inc. Boston MA.

August 28th. The 2nd annual Gaggle meeting.

August 6th, Santa Fe, NM. The 2nd annual quantitative biology meeting (Q-bio). Invited talk. This meeting (run by Los Alamos' Center for Non-Linear Studies)

July 22, Seattle, WA. RosettaCon.

July 17-20, Toronto. ISMB. (ISMB, along with Recomb, is one of the highest profile computational biology meetings). We were invited to give two talks this year, one Highlights talk (highlighting a different paper than presented the year before) and an Invited talk (detailing one of our newer results, the Human proteome folding project).

June 15-21, 2008 . 6th Pathways, Networks and Systems, Chania, Crete, Greece.

May 28th. Department of Energy Knowledge-base workshop. April. 2nd Bertinoro systems biology workshop. Invited Talk. Italy.

April 9. Cold Spring Harbor Bioinformatics Seminar.

March 10. NSF Workshop on Knowledge Management and Visualization.

Jan 20th. 7th Annual McGill Barbados Workshop at the Bellairs Research Station in Holetown, Barbados.

Jan 10th meeting of the Bioinformatics Discussion Group of the New York Academy of Sciences.

## **2007:**

November 26. The Institute of Pure and Applied Mathematics: Mathematics of Knowledge and Search Engines: Workshop IV: Search and Knowledge Building for Biological Datasets.

October, GSK Oncology Division. Invited Seminar.

October, **Department of Energy**, The Office of Science Advanced Scientific Computing Advisory Committee (ASCAC) and Biological and Environmental Research Advisory Committee (BERAC) Panel Meeting. This meeting was meant to assess the results of the first 5 years of GTL funding and to advise on further funding directions.

August, The annual meeting of the New York Structural Biology Group.

July, Vienna, ISMB/ECCB. Highlights Track. This was the joint meeting of the major European and north American computational biology meetings. (Attendance tops 7000). We were given the first talk on Day 1, the highlights talks were also given more time than the normal talks. This was good news for our group.

June, The 2007 Bacillus interest Group annual meeting.

July, Vienna. ISMB. Microbiology Special Interest Group.

August, The annual Gaggle developers Conference (co-organizer).

May, The 2007 Interface meeting. This computational statistics society is devoted to problems on the interface between algorithmic/CS and statistics/modeling.

April: DIMACS Workshop on Discrete Mathematical Problems in Computational Biomedicine

## **2006:**

Keystone Symposium on Systems Biology, Taos, NM, 2006

Rice, Theoretical and Computational Biology Symposium, Houston TX

Prime meeting on Pathway, Databases and modeling tools, Princeton.

Orfeome, 2006, Dana Farber Cancer Center, Harvard Medical School

Princeton, Computer Science Dept Seminar Series.

NYAS, Computational Biology Series, one of three invited speakers for this special event.

UMDNJ-New Jersey Medical School, Bioinformatics Seminar Series

NYU Center for Comparative Functional Genomics Symposium

The 14th World Congress of Psychiatric Genetics, Sardinia

### **Prior to NYU:**

Transcriptome 2005. Shanghai, China. 2005.

2<sup>nd</sup> International E. coli Alliance Conference on Systems Biology. Banff, Canada, 2004.

Invited Speaker. Biomedical Engineering Society Annual Fall meeting. Nashville, TN 2003

IBC workshop: Advances in Structure-Guided Drug Discovery, In silico techniques and SAR approaches, La Jolla, CA, November 15, 2002.

Institute for Complex Adaptive Matter Workshop, "Self-Organizing Biomolecules – the evolving picture", Santa Fe, NM, USA, January 2002.

The 39<sup>th</sup> meeting of the Biophysical Society of Japan, Osaka, Japan, October 2001.

4<sup>th</sup> Critical Assessment of Protein Structure Prediction, Asilomar, CA, USA, December 2000.

### **Mentored and co-mentored students:**

#### **Former:**

Student name: Aviv Madar

Degree program: PhD, Biology

Role (thesis advisor, committee member, examiner): **Advisor**,

Starting Second Year, grad date 2010

Student name: Chris Poultney

Degree program: PhD, CS

Role (thesis advisor, committee member, examiner):

*Co-advising, Denis Shasha is Chris' primary advisor.* grad date 2010

#### **Current:**

Student name: Peter Waltmen

Degree program: PhD, Computer Science/ COB

Role (thesis advisor, committee member, examiner): **Advisor**, Starting Second Year, expected grad date june 2011

Student name: Thadeous Kacmarczyk

Degree program: PhD, Biology

Role (thesis advisor, committee member, examiner): **Advisor**, Starting Second Year, expected grad date july 2010

Student name: Nathan Poslusney

Degree program: PhD, Biology

Role: **Advisor**, Starting Second Year, co-advised with Mike Purugganan

Student name: Timothy Craven

Degree program: PhD, Biology

Role: **Advisor**, Starting First Year

Student name: Alex Greenfield

Degree Program: PhD , COB

Role: **Advisor (w/ H Ostrer, Director Human Genetics Program, NYU-MED)** Starting 3rd year. Expected grad date 2012

Student name: Kevin Drew

Degree Program: PhD , Sackler, Structural Biology Program

Role: **Advisor**, Starting 2nd year.

Student name: Francis Kirigin

Degree Program: PhD , Sackler

Role: **Co-Advisor, w/ Dan Littman**, Starting 1st year.

**Masters students:**

Student Name: Shaila Musharoff

Degree program: Masters, CS

**Former Advisor**, Two semesters of research credits

Student name: Thesis committee: Joe Sarro

Degree program: Masters, Biology

Role (chair, member): **Former Advisor** for internship.

**other PhD advisement:**

Student Name: Sujash Chatterjee

Degree Program: Biology, PhD

Role: committee member, graduated 2010

Student name: Thesis committee: Peter McKenney

Degree program: PhD, biology

Role (chair, member): member, 2<sup>nd</sup> year

Student name: Thesis committee: Jerry Huang

Degree program: PhD, Biology

Role (chair, member): member, 3<sup>rd</sup> year

Student name: Thesis committee: Jeffery Rosenfeld

Degree program: PhD, Biology

Role (chair, member): member, 2<sup>nd</sup> year

Student name: Thesis committee: Kerry Donny-Clark

Degree program: PhD, Biology

Role (chair, member): member, 2<sup>nd</sup> year

**Outreach Experience:**

**NYU-current teaching responsibilities are described in detail below, and in the attached teaching statement.**

**High-school Curriculum Development. 2002-2004.** As part of ISB's ongoing commitment to revolutionizing high-school biology curriculum I have participated in our summer program aimed at creating an inquiry based high-school systems-biology curriculum by working with high school interns and local teachers.

**High School Teacher. 1997-1998. NOVA High, Seattle.** Designed and taught a high school biochemistry curriculum for students with little or no science education ages 14-22.

**Board Member, FSU Center for Participant Education. 1994-1996.** CPE is the oldest free university in the country. Adult education programs, free classes and lectures covering diverse topics were offered free of charge to the community. Taught "Grinding on the Green".

**Teaching and Service at NYU:**

**I aim to provide practical CS training for biologists and hands on experiences with biology data and biology questions for CS majors. Biology is one of the more interdisciplinary fields and therefore presents curriculum challenges that we are a long way from solving. I hope to continue being a part of the ever-necessary renewal of our CS and biology curriculum.**

Both of my classes are described in much greater detail each courses Wiki / Webpage:  
[http://err.bio.nyu.edu/courses/index.php/Main\\_Page](http://err.bio.nyu.edu/courses/index.php/Main_Page) [user: student, passwd: letmein ]

### **G23.1127 / G22.2520-001 Bioinformatics and Genomes (Spring 2007, 2008, 2009 - Bonneau, sole instructor)**

Graduate Course: Bioinformatics. Topics include sequence alignment, motif finding, phylogeny and RNA-structure prediction. Cross listed with CS and Biology. We were successful in forming groups of CS and Bio students and two of the class projects resulted in results that contributed to publications.  
<http://homepages.nyu.edu/~rb133/teaching/biogenomes.html>

### **V22.0480-003 : Special Topics: Computing with Large Data Sets (Computer Science Course, Spring 2009, Richard Bonneau, sole instructor, cross listed with Biology in 2010)**

Enormous collections of data in multiple fields of science and engineering are being gathered that fundamentally challenge the way we analyst large datasets, and thereby our current offerings for undergraduates majoring in CS at NYU. This course discusses some of the associated unprecedented computational challenges, focusing on very large data sets arising in computational biology. High-level languages for mathematical modeling and statistical analysis offer a double-edged sword: use these languages correctly and you'll be able to prototype methods for data analysis and discovery that amaze your co-workers and can be translated into stand-alone code and Web services; but use these language incorrectly and you will end up with inefficient code that is impossible for others to understand. The course is intended to address some of the needed general principles by using the R statistical programming language to analyze large genomic data sets, which provide examples of data types common to several problems.  
[http://err.bio.nyu.edu/courses/index.php/Main\\_Page](http://err.bio.nyu.edu/courses/index.php/Main_Page)

### **Team Taught:**

BIO CORE II: GENES, SYSTEMS & EVOLUTION (G23.1002.001.SP07)

BIO CORE IV: GENES, SYSTEMS & EVOLUTION (G23.2004.001.SP07)

My unit covered Genomics technologies, such as microarray manufacture, error/quality control, proteomics, sequencing. We also covered analysis for these datatypes.

### **Guest Lectures in the following courses:**

(1 Lecture / y ) Microbiology G23.1027

(1 Lecture / y) Genomics and Bioinformtics V23.0038

(1 Lecture / y) Principles of Structural Biology (Sackler, Medical School)

**Tutorials (NYU-medical school PhD students):** These are one-on-one tutorials, given by professors with experience in the area of interest. These tutorials serve the purpose of an alternate proposal. Keren Klein (reviewing papers on protein design), Ben Bartelle (reviewing synthetic biology)

### **Outreach and Undergraduate Research:**

My main outreach activity has been mentoring High-school students over the summer. We have had one High School student in the lab each year, several (Kevin, Devorah, and now Michelle) returned for a second year (3 of the 5 summers we have had overlapping high school students, this helps them fit in). Three of these interns, including the two high school students, made significant contributions and are included on one or more papers. Both of these students were great and went on to MIT and Princeton respectively.

Devorah Kengmana (Stuyvesant High School, NY, Now at MIT)

Kevin Belasco (Scarsdale High School, CT, Now at Princeton)

Alex Pine (CS undergraduate, NYU, 2nd year in lab)

Elena Helman (Biology, Undergraduate, Brown)

Kiran Mayi (Bioengineering, Brooklyn Polytechnic)

### **Committees and University and Departmental Service:**

**Faculty Search Committee** (for both CS and genomics/biology searches)

Annual estimated hours: 90 hrs (much of this time is devoted to the biology search where I am a full member of the committee) Annual estimated hours 100 hrs for both departments.

**Ph.D. Student Recruitment Committee (non-member full participant for both CS and Biology departments, I also help interview Sakler/Skirball applicants to the general and structural biology PhD programs)** (Spring 2006, 2007, 2008). Annual estimated hours: 80 hrs

**NYU Science Faculty Advisory Group on Technology**

**NYU Library liaison committee**

**NYU Academic Standards Committee**

**NYU Innovation Venture Fund's Life Science Investment Review Board Member**

**Biology Computing infrastructure:** I, along with Kris Gunsalus, am currently charged with reorganization of the Biology compute cluster, and the hiring and management of a systems administrator for the Biology compute resources, and better interfacing with NYU's larger IT infrastructure.

**Broad participation in the IT community at NYU,** Author for Connect Magazine (NYU's IT quarterly publication), spoke at ITS Strategic Planning forum (2007), attended and spoke at NYU-Abu Dhabi HPC conference and workshop (2011).

## **Scholarly Activities:**

### **Professional societies:**

ISMB/ ISCB: Intelligent Systems for Molecular Biology / International Society for Computational Biology

ASM: American Society for Microbiology.

AAAS: American Association for the Advancement of Science

NYAS: New York Academy of Science

Interface/Informs: a membership society of computational scientists, statisticians, mathematicians and individuals from related discipline areas interested in the interface between computing science and statistics

### **Editorial and Peer review activities:**

Associate Editor, PLoS Computational Biology

Associate Editor, BMC Systems Biology

### **Reviewer for the following journals:**

Science

Cell

Neuron

Bioinformatics

Nature Molecular Systems Biology

Nature Methods

Nature

Genome Biology

Genome Research

Proteins: Structure, Function, and Bioinformatics

Journal of the American Chemical Society

Biochemistry

Journal of Chemical Physics  
BMC Bioinformatics  
BMC Systems Biology  
PLoS Biology  
PLoS Computational Biology  
PLoS One  
IET Systems Biology  
IEEE Transactions on Signal Processing  
Journal of Computational Chemistry  
Journal of Proteome Research  
DNA and Cell Biology  
Journal of the Royal Society Interface  
Biopolymers

### **Conference organization and panels:**

2011 Co-organizer of 4<sup>th</sup> Berlin Summer Meeting on Computational and Experimental Molecular Biology, Berlin, June 23-25  
2011 Session co-chair, Cold Spring Harbor Meeting: The Biology of Genomes, May 10-14  
RosettaCon, the 2010 Rosetta developers meeting, Main Organizer and organizer for PLoS special collection associated with conference.  
2010 RECOMB DREAM-SB-RG conference Panel member  
2010 ECCB Panel member  
2009 ISMB Panel member  
2009 Co-organizer for Annual Symposium of the Center for Genomics and Systems Biology, NYU

### **Grant Review Panels:**

Reviewer for NSF Interagency Program in Metabolic Engineering (NSF, 2009)  
NSF: Computer and Information Science and Engineering Directorate,  
Intelligent Information Systems Division (2007-current)  
NSF: Physiological and Structural Systems (2007)  
DoE: INCITE (Innovation and Novel Computational Impact on Theory) (2008)  
I have also reviewed by mail (ad hoc) several NIH, NSF and DoE proposals.

### **Collaborations:**

Below is a partial list of people with whom I have collaborated over the last five years:

- |                          |  |
|--------------------------|--|
| 1. Ruedi Aebersold       | ETH, Zurich, Institute of Molecular Systems Biology      |
| 2. Alan Aderem           | ISB, Seattle   |
| 3. David Agus            | USC & Cedars-Sinai Medical Center                        |
| 4. Parker Antin          | University of Arizona, Dept. of Cell Biology and Anatomy |
| 5. Paramjit Arora        | NYU, Dept. of Chemistry                                  |
| 6. David Baker           | U. Washington, Genome Science                            |
| 7. Nitin Baliga          | ISB, Seattle   |
| 8. Mike Boxem            | Harvard Medical School,                                  |
| 9. Tricia Davis          | UW, Genome Science                                       |
| 10. Rob DeSalle          | American Museum of Natural History                       |
| 11. Ken Dill             | Stoney Brook University, Laufer Center                   |
| 12. Michael Dustin       | NYU-medical School                                       |
| 13. Patrick Eichenberger | NYU, Dept. of Biology                                    |
| 14. Marc Facciotti       | UC-Davis   |
| 15. David Goodlett       | UW, Dept. of Medicinal Chemistry                         |
| 16. Kris Gunsalus        | NYU, Dept. of Biology                                    |
| 17. Leroy Hood           | ISB, Seattle   |



18. Daniel Hillis Applied Minds, LLC.
19. Stefan Kempa The Berlin Institute for Medical Systems Biology, MDC
20. Ken Kosik UC Santa Barbara, Neuroscience
21. Brian Kuhlman UNC, Chapel Hill, Dept. of Biochemistry
22. Kent Kirshenbaum NYU, Dept. of Chemistry
23. Josh Labaer Arizona State Biodesign Institute
24. Dan Littman NYU-Medical School
25. Parag Mallick Stanford, [Radiology - Diagnostic Radiology](#)
26. Rick Myers Hudson Alpha
27. Gary Nolan Stanford, Microbiology and Immunology
28. Vistor Ng National Yang-Ming University, Taipei
29. Harry Ostrer NYU Medical School
30. Fabio Piano NYU, Dept. of Biology
31. Michael Purugganan NYU, Dept. of Biology
32. Dallas Rabenstein UC Riverside, Chemistry
33. David Reiss ISB, Seattle
34. Ingo Ruczinski Johns Hopkins, Dept. of Biostatistics
35. Charlie EM Strauss Los Alamos Natl. Labs, Biosciences Division
36. Roland Strong Fred Hutchinson Cancer Research Center
37. Eric Vanden-Eijnden Courant Institute of Mathematical Sciences