William R. Cannon
Senior Scientist and Deputy Division Directory
Biological Sciences Division
Pacific Northwest National Laboratory
Richland WA 99352
Work 509-375-6732
William.cannon@pnnl.gov
wrcannon@gmail.com

Education

Bachelor of Arts	1985	University of California, Santa Cruz in Chemistry
Doctorate	1994	University of Houston, in Biochemistry and Biophysics, Advisors: J. A. McCammon, B. M. Pettitt
Post-doctoral	1998	Pennsylvania State University, Enzymology Advisors: S. J. Benkovic and B. J. Garrison

Research Interests:

Combining modeling & simulation with data analysis, we can understand how cells change as a function of the environmental perturbations, whether the environment is the human gut perturbed by exogenous chemicals or a biofuel reactor that needs to be optimized to produce target chemicals in abundance. On a more ambitious scale, we can understand the principles of self-organization, especially with regard to scaling in space and time and the emergence of biological function and mechanisms.

*Simulation and Modeling***:** We have developed advanced simulation technology that will bring much higher predictive modeling to biology. Briefly, statistical thermodynamics can be used to formulate the law of mass action so that biological processes can be modeled more faithfully and without the need for rate constants. This approach goes beyond the modeling of flux using constraint-based methods and provides information on pathway thermodynamics and prediction of metabolite concentrations. This will allow us to quantitatively model the energetics and dynamics of metabolism and cellular processes.

Data Analysis and Modeling: Models provide the prior knowledge needed to interpret large-scale proteomics, metabolomics and gene expression data. Unfortunately, our ability to effectively utilize data is often limited due to inherent noise in the data, lack of statistical confidence and the lack of models for integrating and making biological sense of the data. My interests are in the use of statistical methods to analyze the data and the use of models as prior knowledge in making sense of the data.

Research Areas: Computational biophysics, biochemistry and proteomics; Modeling and simulation including deterministic and stochastic simulation of metabolism; simulations of state; microbial metabolism; statistics, statistical mechanics and statistical proteomics data analysis; Cloud computing and high performance computing.

Honors and Awards

2013 Computational and Information Sciences Directorate Outstanding Performance Award
2007 Computational and Information Sciences Directorate Outstanding Performance Award
2003 FSD Outstanding Performance Award
2004 FSD Outstanding Performance Award
2000 Monsanto Outstanding Performance Award
1995 NIH Post-doctoral National Research Service Award in Biophysical Chemistry
1992 NIH Pre-doctoral National Research Service Award in Biophysics
1990 Welch Foundation Graduate Award 1990, University of Houston
1985 Honors in Chemistry, University of California

Positions and Employment

2013-present	Deputy Division Director, Biological Sciences Division, PNNL, Richland WA
2010-present	Senior Staff Scientist, Pacific Northwest National Laboratory, Richland, WA
2008-2011	Adjunct Faculty, Washington State University, Pullman, WA
2000-2010	Senior Research Scientist, Pacific Northwest National Laboratory, Richland, WA
1998-2000	Computational Scientist-Genetic Circuit Inference, Monsanto, St. Louis, MO
1994-1998	Post-doctoral Fellow in Enzymology with S.J. Benkovic, Department of Chemistry, Pennsylvania State University, University Park, PA
1993-1994	Instructor, Macromolecular Design, Baylor College of Medicine, Houston, TX
1988	Protein Chemist, University of Washington, Seattle, WA
1988	Chromatographic System Design, Texaco Research Internship, Houston, TX
1987	National Park Ranger, Mt. Rainier National Park, WA
1986-1987	Pharmaceutical/Analytical Chemist, Syntex Corporate Research, Palo Alto, CA
1985	National Park Ranger, Lake Clark National Park, AK
1984	National Park Ranger, Rocky Mountain National Park, CO
1981-1983	National Park Ranger, Mt. Rainier National Park, WA

Professional Activities

- Editorial Board, Nature Scientific Reports.
- Biophysical Society
- American Chemical Society
- American Society for Mass Spectrometry
- American Physical Society

- NIH Study Section ZRG1 BST-F Software and Bioinformatics Fall 2008
- NIH Study Section ZRG1 BST-F Shared Instrumentation, Fall 2009
- Professional Membership (past and current): American Society for Mass Spectrometry, American Chemical Society, Biophysical Society, International Society for Computational Biology
- Reviewer for: Bioinformatics, Biochemistry, Journal of Bioinformatics and Computational Biology, Journal of Proteome Research, Journal of Physical Chemistry, Analytical Chemistry.
- Session Organizer, Pacific Symposium on Biocomputing, 2006
- Session Organizer, Pacific Symposium on Biocomputing, 2007
- Program Committee, Conference on Computational Systems Bioinformatics, Stanford, 2009

Peer reviewer for: *Biochemistry; Nature Biotechnology; Nature Methods; Science; Biophysical Journal; PLoS Computational Biology; Bioinformatics; Journal of Computational Biology; Journal of Physical Chemistry; Journal of the American Chemical Society; Proteins; Journal of Proteome Research; Entropy; Journal of Biological Chemistry; Journal of Molecular Biology*.

Patents

Jarman, K. H., **Cannon, W. R**., Jarman, K. D., Heredia-Langner, A., Peptide identification, U. S. Patent 7,979,214, July 2011.

Technical Publications (in chronological order)

1. **Cannon WR**, JD Madura, RA Thummel, and JA McCammon. (1993) Host-guest binding: Effect of rotational isomers. *J. Amer. Chem. Soc*. 115:879.
2. **Cannon WR**, BM Pettitt, and JA McCammon. (1994) Sulfate anion in water: Model structural, thermodynamic and dynamic properties. *J. Phys. Chem*. 98:6225.

3. **Cannon WR**, JM Briggs, J Shen, JA McCammon, and FA Quiocho. (1995) Conservative and nonconservative mutations in proteins: Anomalous mutations in a transport receptor analyzed by free energy and quantum chemical calculations. *Protein Sci*. 4:387-393.
4. **Cannon WR**, SF Singleton, and SJ Benkovic. (1996) A perspective on biological catalysis. *Nature Structural Biology* 3:821-833.
5. **Cannon WR**, BJ Garrison, and SJ Benkovic. (1997) Electrostatic characterization of enzyme complexes: Evaluation of the mechanism of catalysis of dihydrofolate reductase. *J. Am. Chem. Soc*. 119:2386-2395.
6. **Cannon WR**, BJ Garrison, and SJ Benkovic. (1997) Evaluation of the pH-dependent inhibition of dihydrofolate reductase by methotrexate. *J. Molec. Biol*. 271:656-668.
7. **Cannon WR,** and SJ Benkovic. (1998) Solvation, Reorganization Energy and Biological Catalysis. *J. Biol. Chem*. 273:26257-26260.
8. **Cannon WR**. (2003) Whipple's Disease, Genomics and Drug Therapy, *The Lancet* 361(May 31):1916.
9. **Cannon WR** and KD Jarman. (2003) Improved peptide sequencing using isotope information inherent in tandem mass spectra, *Rapid Commun. Mass Spectrom*. 17:1793-1801.
10. Jarman KD, **WR Cannon**, KH Jarman and A Heredia-Langner. (2003) A Model of Random Sequences for de novo Peptide Sequencing, *Proceedings of the Third IEEE Symposium on Bioinformatics and Bioengineering* 206-213.
11. Heredia-Langner A, **WR Cannon**, KD Jarman and KH Jarman. (2004) Sequence Optimization as an Alternative to de novo Analysis of MS/MS Data, *Bioinformatics,* 20(14): p. 2296-2304.
12. Malard JM, A Heredia-Langner, DJ Baxter, KH Jarman, and **WR Cannon**. (2004) Constrained De Novo Peptide Identification via Multi-objective Optimization, *Proceedings of the Third IEEE International Workshop on High Performance Computational Biology (HiComb)*.
13. Heredia-Langner, A., **WR Cannon**, KD Jarman, and KH Jarman. (2004) De Novo Analysis of Tandem Mass Spectrometry Data as a Non-Deterministic Optimization Problem, *Proceedings of the 2004 International MultiConference in Computer Science and Computer Engineering*.
14. Malard JM, A Heredia-Langner, DJ Baxter, KH Jarman, and **WR Cannon.** (2005) Peptide identification via constrained multi-objective optimization: Pareto-based genetic algorithms. *Concurrency and Computation: Practice and Experience* 17:1687-1704
15. **Cannon WR**, KH Jarman, BM Webb-Robertson, DJ Baxter, CS Oehmen, KD Jarman, A Heredia-Langner, GA Anderson, and KJ Auberry. (2005) A Comparison of Probability and Likelihood Models for Peptide Identification from Tandem Mass Spectrometry Data, *J. Proteome Res.* 4(5):1687-1698
16. Sharp, JL, KK Anderson, DS Daly, DL Auberry, JJ Borkowski, and **WR Cannon.** (2006) Inferring protein associations using protein pull-down assays. 2006 *Proceedings of the American Statistical Association.*
17. Webb-Robertson BJ, **WR Cannon**, JN Adkins. Proteomic Challenges from Molecules to Systems Biology, *Pac. Symp. Biocomput.* 2006, 212-218. (Review)
18. Jarman KD, **WR Cannon**, KH Jarman, and A Heredia-Langner. A Model of Random Sequences for de novo Peptide Sequencing, *Intl. J. Artificial Intel. Tools*, In press.
19. **Cannon WR**, and BM Webb-Robertson. (2007) Computational Proteomics: High-throughput Analysis for Systems Biology. *Pac. Symp. Biocomput*., 403-408. (Review)
20. **Cannon WR,** Taasevigen DJ, Baxter DJ, and Laskin J. (2007) Evaluation of the Influence of Amino Acid Composition in Collision-Induced Fragmentation of Model Peptides. *J. Am. Soc. Mass Spectrom.* 18(9): 1625-1637

21. Sharp JL, KK Anderson, DS Daly, DA Pelletier, **WR Cannon**, DL Auberry, AM White, GB Hurst, DD Schmoyer, WH McDonald, BS Hooker, KD Victry, SH Wiley, MV Buchanan, and V Kery. (2007) Inferring protein-protein associations with Affinity Isolation LC-MS/MS Assays. *J. Prot. Res*., 6 (9), 3788 -3795.

22. Webb-Robertson BM, and **WR Cannon**. Current Trends in Computational Inference from Mass Spectrometry-based Proteomics. *Brief. Bioinform.* 2007 8: 304-317 doi:10.1093/bib/bbm023.

23. Samatova, N. F., Gorin, A., Karpinets, T., Park, B.-H., Pan, C., Straatsma, T. P., Resat, H., **Cannon, W. R**., "Data Driven Computing for Biological Systems", *SciDac Review*, Fall 2007, 10-25. (project review article for DOE periodical on scientific computing).

24. Webb-Robertson BM, **WR Cannon** and CS Oehmen. (2007) Support vector machine classification of probability models and peptide features for improved peptide identification from shotgun proteomics. *Proceedings of the Sixth International Conference on Machine Learning and Applications (ICMLA '07)*, 500-505.

25. **Cannon, W. R.**, B.-J. Webb-Robertson, A. R. Willse[*], M. Singhal[1], L. A. McCue, J. E. McDermott, R. C. Taylor, K. M. Waters, and C. S. Oehmen, An Integrative Computational Framework for Hypotheses-Driven Systems Biology Research in Proteomics and Genomics, in Computational and Systems Biology: Applications and Methods (2008).

26. Taylor RC, M Singhal, DS Daly, KO Domico, AM White, DL Auberry, KJ Auberry, BS Hooker, GB Hurst, JE McDermott, WH McDonald, DA Pelletier, DD Schmoyer, and **WR Cannon**. SEBINI-CABIN: an analysis pipeline for biological network inference, with a case study in protein-protein interaction network reconstruction. *Int. J. Data Mining Bioinform.* 2009.

27. Webb-Robertson, B.J., **Cannon, W. R.**, Oehmen, C. S., Shah, A. R., Gurumoorthi, V., Lipton, M. S., Waters, K. M., *A Support Vector Machine model for the prediction of proteotypic peptides for accurate mass and time proteomics Bioinformatics*, 2008; 24(13); 1503–1509.

28. Pelletier, D. A., Hurst, G. B., Foote, L. J., Lankford, P. K., McKeown, C. K., Lu, T.-Y., Schmoyer, D. D., Shah, M. B., Hervey IV, W. J., McDonald, W. H., Hooker, B. S., **Cannon, W. R**., Daly, D. S., Gilmore, J. M., Wiley, H. S., Auberry, D. L., Wang, Y., Larimer[1], F. W., Kennel, S. J., Doktycz, M. J., Morrell-Falvey, J. L., Owens, E. T., Buchanan, M. V., A General System for Studying Protein-Protein Interactions in Gram-Negative Bacteria, *J. Prot. Res.*, 2008; 7(8); 3319-3328.

29. Sharp, J. S., Borkowski, J. J., Schmoyer, D., Daly, D. S., Purvine, S., **Cannon, W. R.**, Hurst, G. B., Statistically Appraising Process Quality of Affinity Isolation Experiments, *Computational Statistics and Data Analysis,* 2009; 53, 1720-1726.

30. Oehmen, C. S., and **Cannon, W. R.**, Bringing high-performance computing to the biologist's workbench: approaches, applications, and challenges. *J. Phys.: Conf. Ser.* **125** 012052 (9pp) 2008 doi: 10.1088/1742-6596/125/1/012052

31. Kulkarni, G., Kalyanaraman, A., Baxter, D. and **Cannon, W. R**., A Scalable Parallel Approach for Peptide Identification from Large-scale Mass Spectrometry Data, Proceedings of the 2009 IEEE International Conference on Parallel Processing Workshops, p. 423-430, DOI 10.1109/ICPPW.2009.41.

32. Shah, A.R., Adkins, J. N., Baxter, D. J., **Cannon, W. R**., Chavarria-Miranda, D. G., Choudhury, S., Gorton, I., Gracio, D.K., Halter, T. D., Jaitly, N. D., Johnson, J. R., Kouzes, R. T. Macduff, M. C. Marquez, A., Monroe, M. E., Oehmen, C. S., Pike, W.A., Scherrer, C. Villa, O., Webb-Robertson, B. J., Whitney, P. D., Zuljevic, N., Applications in Data-Intensive Computing in *Advances in Computers*, Volume 78 (2010).

33. **Cannon, W.R**. and Rawlins, M. M., A Physicochemical/Thermodynamic Framework for the Interpretation of Peptide Tandem Mass Spectra. *J. Phys. Chem. C*, **2010**, *114* (12), pp 5360–5366 **DOI:** 10.1021/jp905049d

34. Pelletier, D., **Cannon, W.R**., et al. Protein Interactions in the purple non-sulfur phototrophic bacterium *Rhodopseudomonas palustris*. (in preparation 2009).

35. Taylor RC, M Singhal, DS Daly, JM Gilmore, WR Cannon, KO Domico, AM White, DL Auberry, KJ Auberry, BS Hooker, GB Hurst, JE McDermott, WH McDonald, DA Pelletier, DA Schmoyer, and HS Wiley. 2009. "An analysis pipeline for the inference of protein-protein interaction networks." *International Journal of Data Mining and Bioinformatics* 3(4) (Sp. Iss. SI):409-430.

36. Wu, C., Kalyalaraman, A., Cannon, W., 2010. A scalable parallel algorithm for large-scale protein sequence homology detection, 39th International Conference on Parallel Processing (ICPP 2010), September 13-16, 2010, San Diego, CA.

37. **Cannon, W.R**., Rawlins, M. M., Baxter, D., J., Lipton, M., Callister, S., and Bryant, D. A., Large Improvements in MS/MS Based Peptide Identification Rates using a Hybrid Analysis. *J. Proteome Res.*, 2011, *10* (5), pp 2306–2317, **DOI:** 10.1021/pr101130b.

38. Kalyanaraman, A., Latt, B., Baxter, D. J. and **Cannon, W.R**. A MapReduce Implementation of a Hybrid Spectral Library- Database Search Method for Peptide Identification. Bioinformatics (2011) 27 (21): 3072-3073. doi: 10.1093/bioinformatics/btr523.

39. Wu, C., Kalyalaraman, A., **Cannon, W.R.** , pGraph: Efficient Parallel Construction of Large-Scale Protein Sequence Homology Graphs . IEEE Trans. Par. Dist. Sys. 2012 http://doi.ieeecomputersociety.org/10.1109/TPDS.2012.19.

40. Hugo, A., Baxter, D. J., Kulkarni, G., Kalyanaraman, A., Callister, S. J., and **Cannon, W. R**.; Proteotyping of Microbial Communities by Optimization of Matches from Tandem Mass Spectrometry. BioComputing, 2012, pp 225-234, 10.1142/9789814366496_0022.

41. **Cannon, W.R**., and Baxter, D. J., Analyzing Data for Systems Biology: Working at the Intersection of Thermodynamics and Data Analytics, SciDAC 2011 Proceedings, Denver CO.

42. Peterson E.S., McCue L.A., Schrimpe-Rutledge A.C., Jensen J.L., Walker H., Kobold M.A., Webb S.R., Payne S.H., Ansong C., Adkins J.N., **Cannon W.R**, Webb-Robertson B.J., VESPA: software to facilitate genomic annotation of prokaryotic organisms through integration of proteomic and transcriptomic data. BMC Genomics. 2012 Apr 5;13:131. doi: 10.1186/1471-2164-13-131.

43. **Cannon, W. R.** Simulating metabolism with statistical thermodynamics, *PLoS One,* vol. 9, p. e103582, 2014, DOI: 10.1371/journal.pone.0103582.

44. Song, H.-S., **Cannon, W. R**., Belaiev, A. S. and Konapka, A., Mathematical Modeling of Microbial Community Dynamics: A Methodological Review, *Processes* **2014**, *2*(4), 711-752; DOI:10.3390/pr2040711.

45. **Cannon, W. R.,** Concepts, Challenges and Successes in Modeling Thermodynamics of Metabolism, *Front. Bioeng. Biotechnol.*, 2014, DOI: 10.3389/fbioe.2014.00053**.**

46. Thomas, D. G., Jaramillo-Riveri, S., Baxter, D. J. and **Cannon, W. R.**, Comparison of Optimal Thermodynamic Models of the Tricarboxylic Acid Cycle from Heterotrophs, Cyanobacteria and Green Sulfur Bacteria, *J. Phys. Chem.*, 2014, DOI**:** 10.1021/jp5075913.

47. Clancy, C.E., An, G., **Cannon, W.R.**, Liu, Y., May, E.E., Ortoleva, P., Popel, A. S., Sluka, J. P., Su, J., Vicini, P., Zhou, X., Eckmann, D. M., Multiscale modeling in the clinic: drug design and development, Annals of biomedical engineering, 44(9), 2591-2610, DOI: 10.1007/s10439-016-1563-0.

48. **Cannon, W. R**. and Baker, S. E., A Statistical Thermodynamic Approach to the Rates of Coupled Reactions in Condensed Phases. 2016, Submitted.

Research Support History

NSF Kalyanaraman (PI) 10/01/09 – 09/30/12
National Science Foundation
Efficient Algorithms for Data-Intensive Bio-Computing

Project Goals: develop solutions and algorithmic frameworks for problems that conduct *database search* as the core operation and problems that conduct *large-scale graph analysis*.
Role: Co-PI

DOE-ASCR/BER Cannon (PI) 04/01/09 – 03/30/12
Department of Energy
Hierarchical, Multi-Objective Optimization of Microbial Community Proteomics
Project Goals: Develop novel methods needed to describe the proteins and metagenomic functional processes occurring within unsequenced microbial communities.
Role: Principle Investigator

DOE-BER Fredrickson (PI) 10/01/09 – 09/30/12
Department of Energy
PNNL Foundational Scientific Focus Area: Biological Systems Interactions
Project Goals: Address fundamental scientific issues on microbial interactions investigating how microorganisms interact to carry out, in a coordinated manner, complex biogeochemical processes.
Role: Co-PI

DOE-BER Buchanan (PI) 10/01/02 – 09/30/08
Department of Energy
Genomes to Life Center for Molecular and Cellular Systems: A Research Program for Identification and Characterization of Protein Complexes
Project Goals: Develop methods for determining protein-protein interactions in microbes using affinity pull-downs and mass spectrometry. Derive and implement statistical methods for evaluating associations among proteins.
Role: Bioinformatics lead; Co-Investigator

DOE-ASCR Straatsma (PI) 01/28/06 – 2/29/08
Department of Energy
Data Intensive Computing for Complex Biological Systems
Project Goals: Explore the role that new supercomputing architectures can play in moving the frontiers of computational biology forward. Some overlap with proposed work in that both the proposed work and this project examine peptide identifications. This is a short term project, however.
Role: Co-PI

P01NS035965 Stenzel-Pore (PI) 10/01/03 – 04/01/06
NIH/NSD
Molecular Mechanisms of Ischemia
Project Goals: Provide Bioinformatics support for analysis of microarray and genome data.
Role: Co-investigator

DOE Cannon (PI) 05/01/01 – 04/30/04
Department of Energy
Bioinformatics Tools to define the Proteomic State of the Cell
Project Goals: Development of methods for de novo peptide identification.
Role: Principal Investigator

PNNL/DOE Cannon (PI) 04/30/00 – 09/30/02
Laboratory Directed Research and Development
Bioinformatics Platform for High through-put Proteomic and Genomic Analysis
Project Goals: Constrained Inference of Biological Networks from microarray and proteomics data.
Role: Principal Investigator

DOE-BER Fredrickson (PI) 08/01/01 – 09/01/02
Department of Energy
Environmental Sensing, Metabolic Response, and Regulatory Networks in the Respiratory Versatile Bacterium Shewanella oneidensis MR-1
Project Goals: Study of environmental sensing and molecular responses in Shewanella oneidensis
Role: Bioinformatics analysis; Co-Investigator