

# Amarda Shehu

George Mason University  
Computer Science Department  
Engineering Building, MSN 4A5  
Fairfax, Virginia 22030, USA

Tel: 703-993-4135  
Fax: 703-993-1710  
Email: [amarda@gmu.edu](mailto:amarda@gmu.edu)  
<http://cs.gmu.edu/~ashehu>

---

## EMPLOYMENT

**George Mason University, Fairfax, VA** August 2008 – present  
Assistant Professor, Department of Computer Science  
Affiliate Appointment, Department of Bioinformatics and Computational Biology  
Affiliate Appointment, Department of Bioengineering

## EDUCATION

**Rice University, Houston, TX** December 2004 to July 2008  
Ph.D. in Computer Science  
Thesis Title: “*Molecules in Motion: Computing Structural Flexibility*”  
Thesis Advisor: Professor Lydia E. Kavvaki

**Rice University, Houston, TX** August 2002 to December 2004  
M.S. in Computer Science  
Thesis Title: “*Sampling Biomolecular Conformations with Spatial and Energetic Constraints*”  
Thesis Advisor: Professor Lydia E. Kavvaki

**Clarkson University, Potsdam, NY** January 2000 to May 2002  
B.S. in Computer Science  
*Summa Cum Laude* in two majors: (1) Computer Science and (2) Mathematics  
Honor Thesis Title: “*Structural and Computational Complexity Results on Testing Dimension in Graphs*”  
Honor Thesis Advisor: Professor Christino Tamon

## RESEARCH

My research falls in the areas of AI, computational biology, computational biophysics, and bioinformatics. Its objective is to advance understanding of the sequence-structure-function relationship in biological molecules.

**Main research goals:** Advance fundamental knowledge in computer science in modeling complex systems with physical constraints; in particular, marry current understanding of search in AI with knowledge of protein statistical mechanics and biophysics towards a powerful unified probabilistic search framework for addressing complexity and constraints in protein systems. The framework employs robotics-inspired analogies for modeling and search of high-dimensional spaces and evolutionary search-inspired strategies for exploration of nonlinear solution spaces rich in local minima.

**Biological Problems of Interest:** Structural modeling of macromolecular assemblies, structural and thermodynamic characterization of the protein native state, elucidation of functionally-relevant protein conformational ensembles, modeling of protein flexibility, and modeling of geometrically-constrained motions of loop segments in partially-resolved structures.

**Other Related Problems of Interest:** DNA sequence annotation, motif finding, and antimicrobial peptide design, through a combination of techniques that combine model-based search, evolutionary search, machine learning, molecular dynamics simulations, and wet-lab experiments.

## PUBLICATIONS IN PREPARATION

1. Olson B and **Shehu A.** *Novel Algorithmic Frameworks for Protein Conformational Search*, to be submitted to PLoS Comp Biol 2011.
1. Olson B and **Shehu A.** *An Evolutionary Search Framework for Sampling Protein Conformational Spaces*, Application Note, to be submitted to J Bioinformatics 2011.
2. Veltri D, Bishop B, van Hoek M, and **Shehu A.** *Automatic Recognition of Cathelicidins Among Antimicrobial and Decoy Peptide Sequences*, to be submitted to PLoS One 2011.
3. Olson B, De Jong K, and **Shehu A.** *A New Generation of Evolutionary-inspired Search Algorithms for Molecular Spaces*, to be submitted to Evol Comp 2011.
4. Majul A, Bishop B, and **Shehu A.** *A Molecular Dynamics Study of Snake-derived Cathelicidin-like Peptides*, to be submitted to J Comput Chem 2011.

- Molloy K, Barbara D, and **Shehu A.** *Employing Variable-length Fragments in a Probabilistic Sampling of the Protein Conformational Space*, to be submitted to J Bioinformatics 2011.

## JOURNAL PUBLICATIONS

- Kamath U, Compton J, Islamaj-Dogan R, De Jong K, and **Shehu A.** *An Evolutionary Algorithm Approach for Feature Generation from Sequence Data and its Application to DNA Splice-Site Prediction*, Trans Comp Biol and Bioinf 2011 (under review).
- Towson J, Veltri D, **Shehu A.**, and Kabbani N. *A Structural Model of the Human D2-D2 Dopamine Receptor Dimer Complex*, Biol Bulletin 2011 (under review).
- Kamath U, **Shehu A.**, and De Jong K. *A Two-Stage Evolutionary Approach for Effective Classification of Hypersensitive DNA Sequences*, J Bioinf and Comp Biol, 9(3):399-413, 2011.
- Olson B, Molloy K and **Shehu A.** *In Search of the Protein Native State with a Probabilistic Sampling Approach*, J Bioinf and Comp Biol, 9(3):383-398, 2011.
- Shehu A.** and Olson B. *Guiding the Search for Native-like Protein Conformations with an Ab-initio Tree-based Exploration*, Intl J of Robot Res, 29(8):1106-1127, 2010 (invited).
- Hegler JA, Laetzer J, **Shehu A.**, Clementi C, and Wolynes, PG. *Restriction vs. Guidance: Fragment Assembly and Associative Memory Hamiltonians for Protein Structure Prediction*, Proc Natl Acad Sci USA, 106(36):15302-15307, 2009.
- Shehu A.**, Clementi C, and Kavraki LE. *Multiscale Characterization of Protein Conformational Ensembles*, Proteins: Struct Funct Bioinf, 76(4):837-851, 2009.
- Shehu A.**, Kavraki LE, and Clementi C. *Unfolding the Fold of Cyclic Cysteine-rich Peptides*, Prot Sci 17(3):482-493, 2008.
- Shehu A.**, Clementi C, and Kavraki LE. *Sampling Conformation Space to Model Equilibrium Fluctuations in Proteins*, Algorithmica 48(4):303-327, 2007.
- Shehu A.**, Kavraki LE, and Clementi C. *On the Characterization of Protein Native State Ensembles*, Biophys J 92(5):1503-1511, 2007.
- Shehu A.**, Clementi C, and Kavraki LE. *Modeling Protein Conformational Ensembles: From Missing Loops to Equilibrium Fluctuations*, Proteins: Struct Funct Bioinf 65(1):164-179, 2006.

## BOOK CHAPTERS

- Shehu A.** *Conformational Search for the Protein Native State* in Protein Structure Prediction: Method and Algorithms, Wiley Book Series on Bioinformatics, 2009

## CONFERENCE PUBLICATIONS

- Hashmi I, Akbal B, Haspel N, and **Shehu A.** *Evolutionary-guided Sampling of Protein Complex Structures*, Computational Structural Biology Workshop (CSBW), IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM), Atlanta, GA, 2011 (submitted).
- Akbal B, Hashmi I, **Shehu A.**, and Haspel N. *Refinement of Docked Protein Complex Structures Using Evolutionary Information*, Computational Structural Biology Workshop (CSBW), IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM), Atlanta, GA, 2011 (submitted).
- Olson B and **Shehu A.** *Populating Local Minima in the Protein Conformational Space*, IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM), Atlanta, GA, 2011 (accepted).
- Kamath U, De Jong K, and **Shehu A.** *An Evolutionary-based Approach for Feature Generation: Eukaryotic Promoter Recognition*, IEEE Congress on Evolutionary Computation (IEEE CEC), New Orleans, LA, 2011.
- Olson B, Molloy K, and **Shehu A.** *Enhancing Sampling of the Conformational Space Near the Protein Native State*, Intl. Conference on Bio-inspired Models of Network, Information, and Computing Systems (BIONETICS), Boston, MA, 2010 (*best student paper award*).
- Kamath U, **Shehu A.**, and De Jong K. *Feature and Kernel Evolution for Recognition of Hypersensitive Sites in DNA Sequences*, Intl. Conference on Bio-inspired Models of Network, Information, and Computing Systems (BIONETICS), Boston, MA, 2010.

7. Kamath U, **Shehu A**, and De Jong K. *Using Evolutionary Computation to Improve SVM Classification*, IEEE World Congress on Computational Intelligence (WCCI), Barcelona, Spain, 2010.
8. Kamath U, De Jong K, and **Shehu A**. *Selecting Predictive Features for Recognition of Hypersensitive Sites of Regulatory Genomic Sequences with an Evolutionary Algorithm*, Genet and Evol Comp Conf (GECCO), Portland, Oregon, 2010, pg. 179-186.
9. Richardson SM, Olson B, Dymond JS, Burns S, Chandrasegaran S, Boeke JD, **Shehu A**, and Bader JS. *Automated Design of Assemblable, Modular, Synthetic Chromosomes*, Lecture Notes in Computer Science: Parallel Processing and Applied Mathematics (PPAM), 2009, vol. 6068, pg. 280-289.
10. **Shehu A**. *An Ab-initio Tree-based Exploration to Enhance Sampling of Low-energy Protein Conformations*, Robotics: Sci & Sys (RSS), 2009, pg. 31-39.

## TECHNICAL REPORTS

Miles C, Olson B, and **Shehu A**. *Geometry-based Computation of Symmetric Homo-oligomeric Protein Complexes*, Technical Report, GMU-CS-TR-2009-2, 2009.

## SELECTED ABSTRACTS

Olson B and Shehu A. 55th Annual Meeting of the Biophysical Society, Baltimore, MD, 2011.  
 Miles C and Shehu A. International Conference on Intelligent Systems for Molecular Biology & European Conference on Computational Biology (ISMB-EECB), Stockholm, Sweden, 2009.  
 Shehu A and Clementi C. Protein Folding Dynamics Gordon Research Conference, Ventura, CA, 2008.  
 Shehu A, Kavraki L, and Clementi, C. National Meeting of the American Chemical Society, New Orleans, LA, 2008.  
 Shehu A, Clementi C, and Kavraki L. Structural Biology Symposium at UTMB, Galveston, TX, 2005–2007.  
 Shehu A, Clementi C, and Kavraki L. Southwest Regional Meeting of the ACS, Houston, TX, 2006.  
 Shehu A, Clementi C, and Kavraki L. KECK Annual Research Conference, Houston, TX, 2004–2006.  
 Shehu A, Kavraki L, and Clementi C. Multiscale Modeling in Soft Matter and Biophysics Workshop, IPAM UCLA, Los Angeles, CA 2005.  
 Shehu A, Clementi C, and Kavraki L. John Hopkins Folding Meeting, Baltimore, MD, 2005.  
 Shehu A, Clementi C, and Kavraki L. Symposium of Institute of Biosciences and Bioengineering, Houston, TX, 2006.  
 Shehu A and Kavraki L. Rice University Corporate Affiliates Meeting, Houston, TX, 2003–2007.

## SELECTED TALKS

*Novel Algorithmic Frameworks for Protein Conformational Search*, Bioengineering Department Seminar, UMD, College Park, MD, February 10, 2012.  
*Probabilistic Search Algorithms to Compute Conformations of the Protein Native State*, Chemistry Department Seminar, GMU, Fairfax, VA, September 29, 2011.  
*A Probabilistic Framework for the Characterization of the Protein Native State*, Computer Science Seminar, Lehigh University, Bethlehem, PA, March 16, 2011.  
*Combining Evolutionary Algorithms with Supervised Learning to Extract Signals from Biological Data*, BioLearn Workshop, BIONETICS, Boston, MA, December 02, 2010.  
*Mapping Conformational Spaces of Protein Molecules*, Computer Science Seminar, University of Massachusetts at Boston, Boston, MA, December 01, 2010.  
*Characterizing Biological Systems at the Molecular Level*, ACE Scholars Research Highlights, George Mason University, Fairfax, VA, October 13, 2010.  
*A Sampling-based Approach to Compute Biologically-active Protein Conformations. Workshop on Motion Planning: From Theory to Practice*, Robotics: Science and Systems, Saragoza, Spain, June 27, 2010.  
*Probabilistic Methods to Compute Biologically-active Protein Conformations*, Computer Science Seminar, College of William and Mary Williamsburg, VA, June 1, 2010.  
*Computational Aspects of Sequence, Structure, and Function in Protein Molecules*, NCBI CBB Seminar, NIH, MD, Jun 18, 2009.  
*Motions and Assembly of Biological Molecules*, ACE Scholars Program Seminar, George Mason University, VA, Mar 17, 2009.

*Geometric Algorithms for Biological Research: Everything is a Puzzle After All*, GRAND Seminar Series, George Mason University, VA, Oct 16, 2008.

*From Atoms to Molecules to Machines: Computing Protein Motions to Elucidate Function*, Bioinformatics Colloquium, George Mason University, VA, Oct 7, 2008.

*A Multiscale Framework for the Characterization of Protein Native States*, National Meeting of the American Chemical Society (ACS), New Orleans, LA, Apr 6, 2008.

*Computing Structural Flexibility in the Protein Native State*, University of Chicago, Chicago, IL, Feb 6, 2008.

*Characterizing Native Flexibility in Proteins*, Fox Chase Cancer Center, Philadelphia, PA, Jan 23, 2008.

*An Integrated Theoretical/Computational Framework to Characterize Biological Systems at the Nanoscale*, Symposium of the Nanobiology Training Program, Houston, TX, Jul 20, 2007.

*A Computational Framework for Capturing Structural and Motion Properties in Proteins*, Houston Conference on Biomedical Engineering Research, Houston, TX, Feb 9, 2006.

*Sampling Biomolecular Conformations with Spatial and Energetic Constraints*, 10-th Annual Structural Biology Symposium at University of Texas Medical Branch, Galveston, TX, May 21, 2005.

## GRANTS

Communication Core Foundations: Algorithmic Foundations Small Grant for "A Unified Computational Framework to Enhance the Ab-initio Sampling of Native-like Protein Conformations," grant No. 1016995, \$449,998, 9/1/2010 - 8/31/2013.

## SELECTED AWARDS

Recipient of Vice President Research and Economic Development Seed Grants, 2008, 2010

Recipient of Volgenau School Bioengineering Seed Grant, 2009

Recipient of Predoctoral Fellowship of NIH Nanobiology Training Program, 2005-2007

Selected for NSF ADVANCE Workshop on Negotiating Ideal Faculty Position, 2006 (accept. ratio 55/730)

## PROGRAM AND ORGANIZING COMMITTEE

Co-editor of CSBW special issue in J. Bioinf. and Comp. Biol. 2012

Co-editor of Computational Structural Biology Workshop (CSBW), November 12-15, 2011.

Co-organizer and co-chair (with Juan Cortes) of IROS 2011 Symposium Robotics Methods in Computational Biology, September 26-28, 2011.

Co-chair of BIONETICS special issue in J. Bioinf. and Comp. Biol. 2011

Organizer and chair of "Evolutionary Computation and Machine Learning in Bioinformatics Workshop" (BioLearn) at BIONETICS, Boston, Dec. 1-3, 2010

Organizing Committee vice-chair of BIONETICS, Boston, Dec. 1-3, 2010

Co-organizer and co-chair of Bioinformatics track of BIONETICS 2010

## REFEREE

**Grant Proposals:** NSF Comput Core Foundations (panel), NSF Macromol., Supramol. and Nano Chemistry program (ad hoc), NSF Mol. and Cell. Biosciences (ad hoc)

**Journals:** PLoS Comp Biol, Proc Natl Acad Sci USA, J Amer Chem Soc, Biophys J, J Comput Chem, J Bioinf & Comp Biol, J Bioinformatics

**Conferences:** Intl Conf on Comp Sys Bioinf (CSB); Intl Workshop on Algo Found of Robotics (WAFR); IEEE/RSJ Intl Conf on Intelligent Robots and Sys (IROS); Intl Conf Bio-inspired Models of Network, Info, and Comput (Bionetics); IEEE Intl Conf on Networked Digital Technologies (NDT)

## PROFESSIONAL ASSOCIATIONS

IEEE, IEEE Robotics and Automation Society, IEEE Women in Engineering  
 ACM, ACM SIGACT, ACM SIGART, ACM SIGSIM, ACM SIGEVO, ACM SIGKDD  
 International Society for Computational Biology (ICSB); Biophysical Society

## MENTOR

Mentor for Chantilly and Thomas Jefferson High School Students, 2010-present  
 CRA-W Mentor for Undergraduate (DREU) Program 2005, 2009  
 Mentor for Academic Challenges and Enrichment Scholar (ACE), 2009-present (mentee received Honorary  
 Mention in 2010 CRA Undergraduate Researchers Award)  
 Mentor for Aspiring Scientist Summer Internship Program (ASSIP), 2009  
 Csters Faculty Mentor 2008-present  
 Mentor for Real World Design Challenge team, 2008-2009

## SELECTED PROFESSIONAL ACTIVITIES

Organizer of Chantilly High School Computer Science Tour, September 2011  
 Organizer of Chantilly High School Bioengineering Tour, June 17, 2011  
 Two-body Problem Panel Member at Grace Hopper Conference, 2009  
 Bioengineering Program Committee, George Mason University, 2008–present  
 KECK Speaker Selection Committee, Rice University, 2007  
 Rice Educational Outreach, 2006–2008  
 Grace Hopper Celebration of Women in Computing, 2006–2007  
 Rice Undergraduate Research Symposium Judge, 2006  
 Rice University CSTERS Women in Computer Science and Engineering, 2005–2008

## COLLABORATORS

C. Clementi, Rice University; J. Kosecka, George Mason University; J. Lee, NIH; K. De Jong, George Mason  
 University; P.G. Wolynes, Rice University; J. Bader, Johns Hopkins University; B. Bishop, George Mason  
 University; M. L. van Hoek, George Mason University; T. Born, George Mason University; N. Haspel,  
 University of Massachusetts at Boston; D. Barbara, George Mason University

## STUDENTS

**Ph.D. Students (6):** Brian Olson (expected 2014), Kevin Molloy (2015), Daniel Veltri (2015), Uday  
 Kamath (2015), Irina Hashmi (2016), Seyed Farid Hendi (2016)  
**M.S. Students (2):** Amr Majul (expected 2012), Christopher Miles (2010, now at UW Seattle)  
**Undergraduate Students (7):** Sameh Saleh (2011-2012), Talhah Zafar (2011-2012), Justin Towson  
 (2010-2010), Beenish Jamil (2009-2010), Jack Compton (2010-2011), Rachael Chung (DREU, UNCC, 2009),  
 Anahita Mostaghim (ASSIP, Univ. of Virginia, 2009)  
**High School students (7):** Catherine Acio, Aman D'Souza, Sruveera Sathi, Trishna Patel, Nakkul Sreeni-  
 vas, Nishita Upadhyaya (Chantilly, 2011-2012), Subeer Talapatra (Thomas Jefferson, 2010)

## STUDENT THESES

Veltri D. *M.S. Thesis: Sequence-based Classification of Antimicrobial Peptides*, George Mason University,  
 November 2011. Committee: Shehu A (advisor), Vaisman I, and Bishop B.  
 Olson B. *M.S. Thesis: Probabilistic Search Algorithms for Protein Structure Prediction*, George Mason  
 University, October 2011. Committee: Shehu A (advisor), Kosecka J, and Lien J-M.  
 Molloy K. *M.S. Thesis: Variable-Length Fragment Assembly in a Probabilistic Protein Structure Prediction  
 Framework*, George Mason University, June 2011. Committee: Shehu A (advisor), Duric Z, and Lien J-M.

**TEACHING**

- CS123 Computing Across the Curriculum** George Mason University  
designing template to propose course as freshman Gen Ed course as part of the “Computing Across the Curriculum Initiative” at George Mason University, in line with the NSF initiative
- CS689 Planning Motions of Robots and Molecules** George Mason University, Spring 2012  
new course designed and added to CS graduate curriculum
- CS444 Introduction to Computational Biology** George Mason University, Spring 2011  
new course designed and added to CS & EE undergraduate curriculum
- CS583 & CS483 Analysis of Algorithms I** George Mason University, Fall 2010, 2008, Spring 2010
- CS499 Bioinformatics and Computational Biology I** George Mason University, Spring 2010, 2009  
designed as precursor of CS689
- CS795 Geometric Algorithms for Bioinformatics** George Mason University, Fall 2009  
designed as precursor of CS444
- COMP70 From Sequence to Structure: A Research-oriented Approach** Rice University, Spring 2006  
co-designed with Ph.D. advisor, Lydia Kavradi for Connexions project<sup>1</sup>

---

<sup>1</sup><http://edition.cnn.com/2007/TECH/11/08/connexions.learning/>.