

YANG SHEN

Department of Electrical and Computer Engineering, Texas A&M University
3128 TAMU, College Station, Texas 77843-3128, USA

Office: WEB-215I e-mail: yshen@tamu.edu Tel: (979) 862-1694 URL: <http://shen-lab.github.io>

Research Interests

- Multimodal Machine Learning and AI for Health: Modeling biological molecules, systems, and data for mechanistic prediction of variant effects and rational design of functional molecules.

Education and Training

- **Massachusetts Institute of Technology**, CSAIL and BE Cambridge, MA
Postdoctoral Associate in Computer Science and Biological Engineering 09/2008 – 12/2011
- **Boston University** Boston, MA
Ph.D. in Systems Engineering 01/2008
- **University of Science and Technology of China** Hefei, China
Bachelor of Engineering in Automation 06/2002

Professional Experience

- **Texas A&M University**, Dept. of Electrical and Computer Engineering College Station, TX
Assistant Professor, Associate Professor 01/2015 – 08/2021, 09/2021 – present
Dept. of Computer Science and Engineering (Affiliated) 01/2022 – present
Dept. of Translational Medical Sciences (Affiliated) 04/2022 – present
Institute of Biosciences and Technology (Affiliated) 04/2022 – present
- **Massachusetts Institute of Technology** Cambridge, MA
Computer Science and Artificial Intelligence Laboratory, Visiting Scientist 08/2022 – 07/2023
Visiting Assistant Professor 06/2012 – 07/2012
- **Toyota Technological Institute at Chicago** Chicago, IL
Research Assistant Professor in Computer Science 01/2012 – 12/2014

Selected Honors and Awards

- Maximizing Researcher's Impact Award (MIRA), National Institute of General Medical Sciences. MIRA for Established Investigators (2024). MIRA for Early Stage Investigators (2017).
- Electrical & Computer Engineering Outstanding Faculty Award, Texas A&M University. 2021
- CAREER Award, National Science Foundation. 2020
- Engineering Grant Genesis Award, Texas A&M University. 2018
- Dean of Engineering Excellence Award (Assistant Professor Level), Texas A&M University. 2018
- 3rd/51 for hard targets in the 7th CAPRI (Critical Assessment of PRredicted Interactions) 2017–19
2nd/26 for hard CAPRI targets in the 3rd CASP (Critical Assessment of Structure Prediction) 2018

Student Recognition

- Undergraduate research scholar Arghamitra Talukder received the 2022 Computer and Information Science and Engineering (CISE) Graduate Fellowship (CSGrad4US) from NSF. 08/2022

- Ph.D. student Yuning You received the Quality Graduate Student Award (2023) and the Chevron Scholarship (2021) from Electrical and Computer Engineering at Texas A&M University.
- Undergraduate research scholar Maxwell Huffman won the First Place in Overall Capstone at the annual Engineering Project Showcase, for the presentation of his thesis “Interconnected Financial Prediction using Time-Series and Network Data”. The 2021 event at Texas A&M University featured more than 150 team projects showcasing the work of over 1,000 engineering students. He also won the First Place in Senior Capstone Design (Electrical and Computer Engineering/ECE) 05/2021
- Ph.D. student Mostafa Karimi received a 2021 Association of Former Students Distinguished Graduate Student Award for Excellence in Research – Doctoral, one of Texas A&M University’s highest levels of recognition. 03/2021
- Undergraduate advisee Paul Crouther received a National Graduate Education for Minorities (GEM) Full Fellowship from the National GEM Consortium. 2017
- Ph.D. student Mostafa Karimi received an NSF Award to Young Professionals Contributing to Smart and Connected Health from IEEE Engineering in Medicine and Biology Society. 2016

Sponsored Projects (Federal)

- “Unraveling Molecular and Systems-Level Mechanisms of Human Disease-Associated Protein Mutations” (2R35GM124952), **Sole Principal Investigator (PI)**, \$1,976,649, National Institute of General Medical Sciences, National Institutes of Health. 2024–29
- “CAREER: Physics-Constrained Modeling of Molecular Texts, Graphs, and Images for Deciphering Protein-Protein Interactions” (CCF-1943008), **Sole Principal Investigator (PI)**, \$500,000, National Science Foundation. 2020–25
- “Unraveling Molecular and Systems-Level Mechanisms of Human Disease-Associated Protein Mutations” (R35GM124952), **Sole Principal Investigator (PI)**, \$1,676,225, National Institute of General Medical Sciences, National Institutes of Health. 2017–24
- “Targeting ESR1 Mutant Breast Cancer” (BC131458P1), PIs: Geoffrey Greene and Sarat Chandarlapaty, Collaborator: Yang Shen, \$54,000 (Shen), Department of Defense Breakthrough Award. 2014–17
- “Dimension Reduction and Optimization Methods for Flexible Refinement of Protein Docking” (CCF-1347865), **Sole PI**, \$174,987 (including \$88,573 after transfer, CCF-1546278, 2015), National Science Foundation. 2013–17
- “Molecular Dynamics Simulations for Exploring Androgen Receptor Antagonism, Drug-Resistant Mutations, and Antagonist Design” (AR-1-2012), **Sole PI**, 3M CPU hours, Department of Energy Argonne Leadership Computing Facility. 2012–13

Sponsored Projects (Internal)

- “A Multidisciplinary Platform to Develop Thermally Stable and Highly Efficient mRNA Vaccines”, PI: Qing Sun, multiple co-PIs including Yang Shen, ~\$1,300,000 (total) including \$76,500 (Shen), Texas A&M University President’s Excellence Fund (X Grant). 2021-24
- “Deep Learning in Cerebral Magnetic Resonance Imaging at 7T for Advanced Diagnostics”, PI: Roderic Pettigrew, co-PIs: Steven Wright and Yang Shen, \$30,000 (total) including \$10,000 (Shen), Texas A&M University President’s Excellence Fund (T3 Program, non-competitive). 2021–22

- “CRISPR Gene Editing for Healthier Foods and Improved Crop Resilience”, PI: Michael Thomson, multiple co-PIs including Yang Shen, \$500,000 (total) including \$26,187 (Shen), Texas A&M University President’s Excellence Fund (X Grant). 2018–20
- “Characterize Functional Motions of Macromolecules by cryo-EM, Statistical and Computational Modeling”, PI: Junjie Zhang, co-PIs: Jianhua Huang and Yang Shen, \$37,000 (total) including \$12,000 (Shen), Texas A&M University President’s Excellence Fund (T3 Program, non-competitive). 2018–20
- “Discover New Targets and Develop Novel Chemicals to Control Mosquitoes and Ticks using a G Protein-coupled Receptor Pipeline, High-throughput Screening and Computational Modeling”, PI: Patricia Pietrantonio, Co-PI: Yang Shen, \$100,000 (total) including \$9,000 (Shen), Texas A&M AgriLife Research Insect Vector Disease Program. 2018–19
- “Discover New Targets and Develop Novel Chemicals to Deter Mosquito Feeding (or Other Blood-Sucking Vectors Feeding) Using a GPCR1 Pipeline and Computational Modelling”, PI: Patricia Pietrantonio, Co-PI: Yang Shen, \$200,000 (total) including \$16,000 (Shen), Texas A&M AgriLife Research Insect Vector Disease Program. 2016–18

Teaching

- **TAMU**
 - ECEN 303 *Random Signals and Systems* Spring 2015, Fall 2016–19,23
 - ECEN 314 *Signals and Systems* Spring 2018, Fall 2021,24
 - ECEN 766 *Algorithms in Structural Bioinformatics* Fall 2015,18, Spring 2016,17,20,22,24
- **TTIC**
 - Lecturer for *Introduction to Machine Learning* (3 lectures), TTI, Nagoya, Japan. 05/2014
 - Guest lecturer for *Intro. to Bioinformatics & Computational Biology* (2 lectures). 03/2012

Professional Service

- Editorial Board, Journal of Biological Systems, World Scientific. 2018–
- Associate Editor, EURASIP Journal on Bioinformatics and Systems Biology. 2015–17
- Proceedings Committee, ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB), 2020–
- Int’l Conf. on Intelligent Systems for Molecular Biology (ISMB) 2014–17,20–
- Proceedings Co-Chair 2015–19 & Area Chair 2023–24, ACM-BCB.
- Co-Organizer, International Workshop on Biomedical Informatics with Optimization and Machine Learning (BOOM). 2016–2020
- Panelist, National Science Foundation. 03/2014, 12/2017, 02/2019,11/2019, 03/2020, 11/2024
- Ad Hoc Reviewer, National Science Foundation. 05/2014, 01/2020, 12/2022

Guest-Edited Journal Special Issues

- Z Wang, V Patel, B Yao, S Jiang, H Lu, **Y Shen** (Guest Editors) (2022), “Learning with Multimodal Data for Biomedical Informatics”, a Special Section of the IEEE Transactions on Circuits and Systems for Video Technology.

- Z Wang, **Y Shen**, S Huang, T McDonald, and J Zhou (Guest Editors) (2018), “Artificial Intelligence in Healthcare Informatics: Learning, Optimization, and Decision Making”, a Special Issue of IISE (Institute of Industrial and Systems Engineers) Transactions on Healthcare Systems Engineering, Taylor and Francis.
- S Huang, **Y Shen**, J Zhou, Z Wang, and Q Ling (Guest Editors) (2017), “Biomedical Informatics with Optimization and Machine Learning”, a Special Issue of EURASIP (European Association for Signal Processing) Journal on Advances in Signal Processing (JASP), Springer Open.
- **Y Shen**, S Huang, J Zhou, Z Wang, and Q Ling (Guest Editors) (2016), “Biomedical Informatics with Optimization and Machine Learning”, a Special Issue of EURASIP Journal on Bioinformatics and Systems Biology (JBSB), Springer Open.
Editorial: S Huang, J Zhou, Z Wang, Q Ling, and **Y Shen** (2017), “Biomedical Informatics with Optimization and Machine Learning”, EURASIP JBSB, 2017:4.

Book Chapter

- D Beglov, R Brenke, G-Y Chuang, D Hall, M Landon, CH Ngan, **Y Shen**, S Thiel, B Zerbe, D Kozakov, and S Vajda, “Identification of Druggable Hot Spots on Proteins and in Protein Protein Interfaces”, in *Computational Protein-Protein Interactions* (Ed. Ruth Nussinov and Gideon Schreiber), CRC Press, June 2009, pages 253–280.

Patent

- MD Balbas, MJ Evans, CL Sawyers, **Y Shen**, D Hosfield, and GL Greene “Modulators of Resistant Androgen Receptor”, US Patent 9,650,359, approved 2017.

Publications

Citations: ~10K; H-Index: 36

Mentee: ^U Undergraduate student ^M Master’s student ^D Doctoral student ^P Postdoctoral researcher
 = Equal contribution * Corresponding author(s)

32–68: *Selected Independent Work*; 12–31: *Other Independent Work*; 1–11: *Earlier Work*.

AI/ML: Theories and Algorithms

68. Y You^D, R Zhou, J Park^P, H Xu^M, C Tian, Z Wang, **Y Shen** (2024), “Latent 3D Graph Diffusion”, *Proceedings of the 12th International Conference on Learning Representations (ICLR 2024)*, **acceptance rate: 31%**.
67. Y You^D, T Chen, Z Wang, **Y Shen** (2023), “Graph Domain Adaptation via Theory-Grounded Spectral Regularization”, *Proceedings of the 11th International Conference on Learning Representations (ICLR 2023)*, **acceptance rate: 31.8%**.
66. T Wei⁼, Y You^{D=}, T Chen, **Y Shen**, J He, Z Wang (2022), “Augmentations in Hypergraph Contrastive Learning: Fabricated and Generative”, *Advances of Neural Information Processing Systems (NeurIPS 2022)*, **acceptance rate: 25.6%**.
65. Y You^D, Y Cao^D, T Chen, Z Wang, **Y Shen** (2022), “Bayesian Modeling and Uncertainty Quantification for Learning to Optimize: What, Why, and How”, *Proceedings of the 10th International Conference on Learning Representations (ICLR 2022)*, **acceptance rate 32.3%**.

64. Y You^{D=}, T Chen⁼, Z Wang, **Y Shen** (2022), “Bringing Your Own View: Graph Contrastive Learning without Prefabricated Data Augmentations”, *Proceedings of the 15th ACM International Conference on Web Search and Data Mining (WSDM 2022)*, pages 1300–1309, **acceptance rate 20.2%. Cited >70 times.**
63. Y You^{D=}, T Chen⁼, **Y Shen***, Z Wang* (2021), “Graph Contrastive Learning Automated”, *Proceedings of the International Conference on Machine Learning (ICML 2021)*, **long talk, acceptance rate: 3.0%, cited >500 times.**
62. Y You^{D=}, T Chen⁼, Y Sui, T Chen, Z Wang, **Y Shen** (2020), “Graph Contrastive Learning with Augmentations”, *Advances of Neural Information Processing Systems (NeurIPS 2020)*, **acceptance rate: 20.1%, cited >2000 times.**
61. Y You^{D=}, T Chen⁼, Z Wang, **Y Shen*** (2020), “When Does Self-Supervision Help Graph Convolutional Networks?”, *Proceedings of the International Conference on Machine Learning (ICML 2020)*, **acceptance rate: 21.8%, cited >240 times.**
60. Y You^{D=}, T Chen⁼, Z Wang, **Y Shen** (2020), “L²-GCN: Layer-Wise and Learned Efficient Training of Graph Convolutional Networks”, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020)*, **acceptance rate: 22.1%, cited >100 times.**
59. Y Cao^D, T Chen, Z Wang, **Y Shen** (2019), “Learning to Optimize in Swarms”, *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, **acceptance rate: 21.6%. Cited >60 times.**

AI/ML: Algorithms for Bioinformatics

58. J Park^P and **Y Shen** (2024), “Equivariant Blurring Diffusion for Hierarchical Molecular Conformer Generation”, *Advances of Neural Information Processing Systems (NeurIPS 2024)*, **acceptance rate: 25.8%.**
57. Y Sun^D and **Y Shen*** (2024), “Structure-Informed Protein Language Models Are Robust Predictors for Variant Effects”, *Human Genetics*, DOI: 10.1007/s00439-024-02695-w.
56. W Tan^D and **Y Shen*** (2023), “Multimodal Learning of Noncoding Variant Effects using Genome Sequence and Chromatin Structure”, *Bioinformatics* 39(9): btad541.
55. Y You^D and **Y Shen*** (2022), “Cross-Modality and Self-Supervised Protein Embedding for Compound–Protein Affinity and Contact Prediction”, *Bioinformatics*, 38(S_2), ii68–ii74. (Also the Proceedings of ECCB 2022, **acceptance rate: 17.4%**)
54. Y Cao^{D*}, P Das*, P-Y Chen*, V Chenthamarakshan*, I Melnyk*, **Y Shen*** (2021), “Fold2Seq: A Joint Sequence(1D)-Fold(3D) Embedding-based Generative Model for Protein Design”, *Proceedings of the International Conference on Machine Learning (ICML 2021)*, **acceptance rate: 21.5%.**
53. Y Cao^D and **Y Shen*** (2021), “TALE: Transformer-based protein function Annotation with joint sequence-Label Embedding”, *Bioinformatics* 37(18), 2825–2833. **Cited >80 times.**
52. M Karimi^{D=}, D Wu^{M=}, Z Wang, **Y Shen*** (2021), “Explainable Deep Relational Networks for Predicting Compound-Protein Affinities and Contacts”, *Journal of Chemical Information and Modeling* 61(1), 46–66. **Cited >50 times.**

51. M Karimi^{D=}, S Zhu^{D=}, Y Cao^{D=}, **Y Shen*** (2020), “De Novo Protein Design for Novel Folds using Guided Conditional Wasserstein Generative Adversarial Networks”, *Journal of Chemical Information and Modeling* 60(12), 5667–81. **Featured cover. Cited >90 times.**
50. Y Cao^D and **Y Shen*** (2020), “Bayesian Active Learning for Optimization and Uncertainty Quantification in Protein Docking”, *Journal of Chemical Theory and Computation* 16(8): 5334–5347.
49. M Karimi^{D=}, A Hasanzadeh⁼, **Y Shen*** (2020), “Network-Principled Deep Generative Models for Designing Drug Combinations as Graph Sets”, *Bioinformatics* 36(S1): i445–i454. (Also the Proceedings of ISMB 2020, **acceptance rate: 20%**)
48. Y Cao^D and **Y Shen*** (2020), “Energy-based Graph Convolutional Networks for Scoring Protein Docking Models”, *Proteins: Structure, Function, and Bioinformatics* 88(8):1091–1099. **Cited >50 times.**
47. Y Cao^D, Y Sun^D, M Karimi^D, H Chen^M, O Moronfoye^U, **Y Shen*** (2019), “Predicting Pathogenicity of Missense Variants with Weakly Supervised Regression”, *Human Mutation* 40(9): 1579–1592.
46. M Karimi^D, D Wu^M, Z Wang, and **Y Shen*** (2019), “DeepAffinity: Interpretable Deep Learning of Compound-Protein Affinity through Unified Recurrent and Convolutional Neural Networks”, *Bioinformatics* 35(18): 3329–3338. **Cited >450 times.**
45. M Karimi^D and **Y Shen*** (2018), “iCFN: an Efficient Exact Algorithm for Multistate Protein Design”, *Bioinformatics*, 34(17), i811–i820. (Also the Proceedings of ECCB 2018, **acceptance rate: 17%**)
44. H Chen^{M=}, Y Sun^{M=}, and **Y Shen*** (2017), “Predicting Protein Conformational Changes for Unbound and Homology Docking: Learning from Intrinsic and Induced Flexibility”, *Proteins: Structure, Function, and Bioinformatics*, 85(3), 544–556.
43. Tomasz Oliwa^P and **Y Shen*** (2015) “cNMA: A Framework of Encounter Complex-based Normal Mode Analysis to Model Conformational Changes in Protein Interactions”, *Bioinformatics* 31(12), i151–i160. (Also the Proceedings of ISMB 2015, **acceptance rate: 17.4%**)
42. **Y Shen*** (2013) “Improved Flexible Refinement of Protein Docking in CAPRI Rounds 22–27”, *Proteins: Structure, Function and Bioinformatics* 81(12), 2129–2136.

AI/ML: Cross-Disciplinary Collaborations

41. S Irani⁼, W Tan^{D=}, Q Li⁼, W Toy, C Jones, M Gadiya, A Marra, JA Katzenellenbogen, KE Carlson, BS Katzenellenbogen, M Karim^D, RS Rajappachetty, IS Del Priore, J Reis-Filho, **Y Shen***, S Chandarlapaty* (2024), “Somatic Estrogen Receptor α Mutations that Induce Dimerization Promote Receptor Activity and Breast Cancer Proliferation”. *Journal of Clinical Investigation* 134(1):e163242.
40. EK Ramos, CF Tsai, Y Jia, Y Cao^D, M Manu, Taftaf R, . . ., T Shi⁼, **Y Shen⁼**, NK Dashzeveg⁼, H Liu⁼ (2022). “Machine Learning-Assisted Elucidation of CD81–CD44 Interactions in Promoting Cancer Stemness and Extracellular Vesicle Integrity”. *elife*, 11:e82669.
39. M Kawaguchi M, N Dashzeveg, Y Cao^D, Y Jia, X Liu*, **Y Shen***, H Liu* (2020), “Extracellular domains I–II of cell-surface glycoprotein CD44 mediate its trans-homophilic dimerization and tumor cluster aggregation”, *Journal of Biological Chemistry* 295, 2640–2649.

38. JC Kang, W Sun, P Khare, M Karimi^D, X Wang, **Y Shen**, RJ Ober, ES Ward (2019), “Engineering a HER2-Specific Antibody-Drug Conjugate to Increase Lysosomal Delivery and Therapeutic Efficacy”, *Nature Biotechnology*, 37, 523–526. **Cited >80 times.**
37. TA Knijnenburg et. al. including **Y Shen**, M Karimi^D, and H Chen^M (2018), “Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas”, *Cell Reports*, 23(1), 239–254. **Among “Best of 2018” selected by Cell Reports. Cited >970 times.**
36. SW Fanning, R Jeselsohn, V Dharmarajan, CG Mayne, M Karimi^D, G Buchwalter, R Houtman, W Toy, CE Fowler, M Laine, KE Carlson, TA Martin, J Nowak, JC Nwachukwu, DJ Hosfield, S Chandarlapaty, E Tajkhorshid, K Nettles, PR Griffin, **Y Shen**, JA Katzenellenbogen, M Brown and GL Greene (2018), “The SERM/SERD Bazedoxifene Disrupts ESR1 Helix 12 to Overcome Acquired Hormone Resistance in Breast Cancer Cells”, *eLife*, 7:e37161. **Cited 100 times.**
35. M Vercruysse, C Köhrer, **Y Shen**, S Proulx, A Ghosal, BW Davies, UL RajBhandary, GC Walker (2016) “Identification of YbeY-Protein Interactions Involved in 16S rRNA Maturation and Stress Regulation in *Escherichia coli*”, *mBio* 7(6), e01785–16
34. SW Fanning, CG Mayne, V Dharmarajan, KE Carlson, TA Martin, SJ Novick, W Toy, B Green, S Panchamukhi, BS Katzenellenbogen, E Tajkhorshid, PR Griffin, **Y Shen**, S Chandarlapaty, JA Katzenellenbogen, GL Greene (2016) “Estrogen Receptor alpha Somatic Mutations Y537S and D538G Confer Breast Cancer Endocrine Resistance by Stabilizing the Activating Function-2 Binding Conformation”, *eLife* 5, e12792. **Cited ~300 times.**
33. W Toy, **Y Shen**, H Won, B Green, RA Sakr, M Will, Z Li, K Gala, S Fanning, TA King, C Hudis, D Chen, T Taran, G Hortobagyi, GL Greene, M Berger, J Baselga, and S Chandarlapaty (2013) “*ESR1* Ligand-Binding Domain Mutations in Hormone-Resistant Breast Cancer”, *Nature Genetics* 45(12), 1439–1445. **Cited >1200 times.**
32. MD Balbas, MJ Evans, DJ Hosfield, J Wongvipat, V Arora, PA Watson, Y Chen, GL Greene, **Y Shen***, and CL Sawyers* (2013) “Overcoming Mutation-Based Resistance to Antiandrogens with Rational Drug Design”, *eLife* 2, e00499. **Cited >450 times.**

Other Independent Work

31. M Zitnik et al., including **Y Shen** (2024), “Current and Future Directions in Network Biology”, *Bioinformatics Advances*, 4(1): vbae099.
30. SL Stenton et al., including **Y Shen**, Y Sun^D, W Tan^D, R Yin^D, and Y You^D (2024), “Critical Assessment of Variant Prioritization Methods for Rare Disease Diagnosis within the Rare Genomes Project, Human Genomics 18(1): 44.
29. The Critical Assessment of Genome Interpretation Consortium including Y Cao^D, H Chen^M, M Karimi^D, O Moronfoye^U, Y Sun^D and **Y Shen** (2024), “CAGI, the Critical Assessment of Genome Interpretation, Establishes Progress and Prospects for Computational Genetic Variant Interpretation Methods”, *Genome Biology* 25: 53.
28. MF Lensink et al., including S Zhu^D, R Yin^D, Y Sun^D, and **Y Shen** (2023), “Impact of AlphaFold on structure prediction of protein complexes: The CASP15–CAPRI experiment”, *Proteins: Structure, Function, and Bioinformatics* 91(12): 1658–1683.

27. J Schreiber et al., including W Tan^D, Y Sun^D, and **Y Shen** (2023), “The ENCODE Imputation Challenge: a critical assessment of methods for cross-cell type imputation of epigenomic profiles”, *Genome Biology* 24: 79.
26. L El-Shennawy et al., including **Y Shen** (2022), “Circulating ACE2-expressing extracellular vesicles block broad strains of SARS-CoV-2”, *Nature Communications* 13: 405.
25. MF Lensink et al., including Y Sun^D, S Zhu^D and **Y Shen** (2021), “Prediction of protein assemblies, the next frontier: The CASP14–CAPRI experiment”, *Proteins: Structure, Function, and Bioinformatics* 89(12): 1800–1823.
24. R Taftaf, X Liu, S Singh, Y Jia, NK Dashzeveg, AD Hoffmann, L El-Shennawy, EK Ramos, V Adorno-Cruz, EJ Schuster, D Scholten, D Patel, Y Zhang, AA Davis, C Reduzzi, Y Cao^D, P D’Amico, **Y Shen**, M Cristofanilli, WA Muller, V Varadan, H Liu (2021), “ICAM1 initiates CTC cluster formation and trans-endothelial migration in lung metastasis of breast cancer”, *Nature Communications* 12, 4867. **Cited >100 times.**
23. A Cichonska, ..., The IDG-DREAM Drug-Kinase Binding Prediction Challenge Consortium (including M Karimi^D, D Wu^M and **Y Shen**), ..., T Aittokallio (2021), “Crowdsourced mapping of unexplored target space of kinase inhibitors”, *Nature Communications* 12, 3307.
22. MF Lensink et al. including Y Cao^D and **Y Shen** (2019), “Blind prediction of homo- and hetero-protein complexes: The CASP13–CAPRI experiment”, *Proteins: Structure, Function, and Bioinformatics* 87(12): 1200–1221.
21. XS Wang, PH Chen, JT Hampton, JM Tharp, CA Reed, SK Das, DS Wang, HS Hayatshahi, **Y Shen**, J Liu, WR Liu (2019), “A Genetically Encoded, Phage-Displayed Cyclic-Peptide Library”, *Angewandte Chemie International Edition* 58(44): 15904–9.
20. A Voskanyan, P Katsonis, O Lichtarge, V Pejaver, P Radivojac, SD Mooney, E Capriotti, Y Bromberg, Y Wang, M Miller, PL Martelli, C Savojardo, G Babbi, R Casadio, Y Cao^D, Y Sun^D, **Y Shen**, A Garg, D Pal, Y Yu, CD Huff, SV Tavtigian, E Young, SL Neuhausen, E Ziv, LR Pal, G Andreoletti, S Brenner, MG Kann (2019), “Assessing the Performance of in-silico Methods for Predicting the Pathogenicity of Variants in the Gene *CHEK2*, among Hispanic Females with Breast Cancer”, *Human Mutation* 40(9): 1612–1622.
19. MS Cline, G Babbi, S Bonache, Y Cao^D, R Casadio, X de la Cruz, O Diez, S Gutierrez-Enriquez, P Katsonis, C Lai, O Lichtarge, PL Martelli, G Mishne, A Moles-Fernandez, G Montalban, SD Mooney, R O’Conner, L Ootes, S Ozkan, N Padilla, KA Pagel, V Pejaver, P Radivojac, C Riera, C Savojardo, **Y Shen**, Y Sun^D, S Topper, MT Parsons, AB Spurdle, DE Goldgar, The ENIGMA Consortium (2019), “Assessment of Blind Predictions of the Clinical Significance of *BRCA1* and *BRCA2* Variants”, *Human Mutation* 40(9): 1546–1556.
18. C Savojardo, M Petrosino, G Babbi, S BoVo, C Corbi-Verge, R Casadio, P Fariselli, L Folkman, A Garg, M Karimi^D, P Katsonis, PM Kim, O Lichtarge, PL Martelli, A Pasquo, D Pal, **Y Shen**, AV Strokach, P Turina, Y Zhou, G Andreoletti, S Brenner, R Chiaraluce, V Consalvi, E Capriotti (2019), “Evaluating the Predictions of the Protein Stability Change upon Single Amino Acid Substitutions for the FXN CAGI5 Challenge”, *Human Mutation* 40(9): 1392–1399.
17. MP Menden et al. including AstraZeneca-Sanger Drug Combination DREAM Consortium (members include J Abante^M, H Chen^M, M Karimi^D, S Xie^D, and **Y Shen**) (2019), “Community assessment

to advance computational prediction of cancer drug combinations in a pharmacogenomic screen”, *Nature communications* 10(1): 1–17. **Mentored students through ECEN 689 course project.**

16. X Liu, R Taftaf, M Kawaguchi, Y-F Chang, W Chen, D Entenberg, Y Zhang, L Gerratana, S Huang, D Patel, E Tsui, V Adorno-Cruz, S Chirieleison, Y Cao^D, A Harney, S Patel, A Patsialou, **Y Shen**, S Avril, H Gilmore, J Lathia, D Abbott, M Cristofanilli, J Condeelis, and H Liu (2019), “Homophilic CD44 Interactions Mediate Tumor Cell Aggregation and Polyclonal Metastasis in Patient-Derived Breast Cancer Models”, *Cancer Discovery*, 9(1), 96–113. **Cited >330 times.**
15. PV Pietrantonio, C Xiong, RJ Nachman, and **Y Shen** (2018), “G Protein-Coupled Receptors in Arthropod Vectors: Omics and Pharmacological Approaches to Elucidate Ligand-Receptor Interactions and Novel Organismal Functions”, *Current Opinions in Insect Science*, 29, 12–20.
14. MF Lensink and about 100 others including **Y Shen** (2016), “Prediction of Homoprotein and Heteroprotein Complexes by Protein Docking and Template-based Modeling: A CASP–CAPRI Experiment”, *Proteins: Structure, Function, and Bioinformatics*, 84(S1), 323–348.
13. CR Drake, L Estévez-Salmerón, P Gascard, **Y Shen**, TD Tlsty, EF Jones (2015) “Towards Aspirin-Inspired Self-Immolating Molecules which Target the Cyclooxygenases”, *Organic & Biomolecular Chemistry* (Royal Society of Chemistry) 13(45): 11078-86.
12. MF Lensink and 57 others including **Y Shen** (2014) “Blind Prediction of Interfacial Water Positions in CAPRI”, *Proteins: Structure, Function and Bioinformatics* 82(4), 620–632.

Earlier Trainee Work

11. **Y Shen**, ML Radhakrishnan, and B Tidor (2015) “Molecular Mechanisms and Design Principles for Promiscuous Inhibitors to Avoid Drug Resistance: Lessons Learned from HIV-1 Protease Inhibition”, *Proteins: Structure, Function and Bioinformatics*, 83(2), 351–372.
10. **Y Shen**, MD Altman, A Ali, MNL Nalam, H Cao, TM Rana, CA Schiffer, and B Tidor (2013) “Testing the Substrate-Envelope Hypothesis with Designed Pairs of Compounds”, *American Chemical Society (ACS) Chemical Biology* 8(11), 2433–2441.
9. **Y Shen**, MK Gilson, and B Tidor (2012) “Charge Optimization Theory for Induced-Fit Ligands”, *Journal of Chemical Theory and Computation* 8(11), 4580–4592.
8. D Kozakov, DR Hall, D Beglov, R Brenke, SR Comeau, **Y Shen**, K Li, J Zheng, P Vakili, ICh Paschalidis, and S Vajda (2010) “Achieving Reliability and High Accuracy in Automated Protein Docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13–19”, *Proteins: Structure, Function and Bioinformatics* 78(15), 3124–3130.
7. **Y Shen**, ICh Paschalidis, P Vakili, and S Vajda (2008) “Protein Docking by the Underestimation of Free Energy Funnels in the Space of Encounter Complexes”, *PLoS Computational Biology* 4(10), e1000191.
6. **Y Shen**, R Brenke, D Kozakov, SR Comeau, D Beglov, and S Vajda (2007) “Docking with PIPER and Refinement with SDU in Rounds 6–11 of CAPRI”, *Proteins: Structure, Function and Bioinformatics* 69(4), 734–742.
5. SR Comeau, D Kozakov, R Brenke, **Y Shen**, D Beglov, and S Vajda (2007) “ClusPro: Performance in CAPRI Rounds 6–11 and the New Server”, *Proteins: Structure, Function and Bioinformatics* 69(4), 781–785.

4. ICh Paschalidis, **Y Shen**, P Vakili, and S Vajda (2007) "SDU: A Semi-Definite Programming-Based Underestimation Method for Stochastic Global Optimization in Protein Docking", *IEEE Transactions on Automatic Control* 52(4), 664–676.
3. **Y Shen**, P Vakili, S Vajda, and ICh Paschalidis (2007) "Optimizing Noisy Funnel-like Functions on the Euclidean Group with Applications to Protein Docking", *Proceedings of the 46th IEEE Conference on Decision and Control (CDC'07)*, 4545–4550, Dec. 12-14, 2007, New Orleans, Louisiana.
2. ICh Paschalidis, **Y Shen**, P Vakili, and S Vajda (2006) "Protein–Protein Docking with Reduced Potentials by Exploiting Multi-Dimensional Energy Funnels", *Proceedings of the 28th IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC'06)*, 5330–5333, Aug. 2006, New York City, New York.
1. ICh Paschalidis, **Y Shen**, P Vakili, and S Vajda (2005) "Semi-Definite Programming-Based Underestimation Method for Global Optimization in Molecular Docking", *Proceedings of the 44th IEEE Conference on Decision and Control (CDC'05)*, 3675–3680, Dec. 2005, Seville, Spain.

Conference Abstracts

• Talks

28. "Interpretable Affinity Prediction and Generative Design for Protein-Ligand Interactions", *8th CAPRI Evaluation Meeting (CAPRI8)*, Feb. 14–16, 2024, Grenoble, France.
27. "Flexible Refinement, Quality Estimation and Uncertain Quantification for Protein–Protein Docking", *American Chemical Society Fall 2023 National Meeting (ACS'23-F)*, Tack Kuntz Symposium, **invited**, Aug. 13–17, 2023, San Francisco, California.
26. "Machine Learning for Protein-Ligand Affinity Prediction: Explainable, Cross-Modal and Self-supervised", *American Chemical Society Fall 2023 (ACS'23-F)*, Symposium "Taking a Deep Dive into Chemical Space", Aug. 13–17, 2023, San Francisco, California.
25. "Explainable Data-Driven Prediction of Protein–Ligand Binding Affinity", **invited**, *10th Conference on Modeling of Protein Interactions (MPI'23)*, May 25–27, 2023, Lawrence, Kansas.
24. "Cross-Modality and Self-Supervised Protein Embedding for Compound-Protein Affinity and Contact Prediction", Y You^D (presenter) and Y Shen, **selected as oral presentation**, *29th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB'22)*, July 10–14, 2022, Madison, USA.
23. "Zero-Shot Variant Effect Prediction with Protein Language Models", Y Sun^D and Y Shen, **selected as oral presentation**, *6th Conference on Critical Assessment of Genome Interpretation (CAGI6)*, May 14–16, 2022, San Francisco, USA.
22. "Transformer-based Protein Function Annotation with Joint Sequence–Label Embedding", Y Cao^D (presenter) and Y Shen, **selected as oral presentation**, *29th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB'21)*, July 25–30, 2021, virtual.
21. "Interpretable Deep Learning for Predicting Compound-Protein Interactions", **invited**, *CNLS (Center for Nonlinear Studies) Virtual Conference – Machine Learning in Chemical and Materials Sciences, Los Alamos National Laboratory*, virtual, May 13, 2021.
20. "Interpretable Deep Learning for Binding Affinity Prediction and Deep Generative Models for Molecular Design", **invited**, *IDDD (Innovative Drug Design and Development) Round Table Workshop: AI in Drug Discovery and Development, Gulf Coast Consortia*, virtual, Feb. 11, 2021.

19. “De novo protein design for novel folds with guided & conditional Wasserstein GAN”, **selected as oral presentation**, *28th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB’20), online, July 13–16, 2020.
18. “Explainable deep models for compound-protein binding affinity prediction and deep generative models for protein design”, *AI TechConnect Virtual Summit (AI for Biomaterials and Drug Design)*, online, June 9, 2020.
17. “Bayesian active learning for optimization and uncertainty quantification in protein docking”, **selected as oral presentation**, *27th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB’19), July 21–25, 2019, Basel, Switzerland. Also presented at *7th CAPRI Evaluation Meeting* (CAPRI7), Hinxton, UK, April 3–5, 2019.
16. “Interpreting System-level Cancer Mechanisms through Rule Learning”, H Chen^{M=}, S-P Deng^{P=}, and Y Shen, *MidSouth Conference on Computational Biology and Bioinformatics (MCBIOS)* 2019, March 30, 2019, Birmingham, Alabama. **Dr. Deng was selected as a podium speaker to give the talk.**
15. “iCFN: an efficient exact algorithm for multistate protein design”, **selected as oral presentation**, *26th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB’18), July 6–10, 2018, Chicago, Illinois.
14. “Predicting missense mutational effects on cancer pathogenicity”, **selected as oral presentation**, *5th Conference on Critical Assessment of Genome Interpretation* (CAGI5), July 5–7, 2018, Chicago, Illinois.
13. “Anticipating cancer mutations through combinatorial protein design”, **invited**, *18th Yale Workshop on Adaptive and Learning Systems*, June 21–23, 2017, Yale University, Connecticut.
12. “Predicting protein conformational changes upon binding and mutation”, **invited**, *8th Conference on Modeling of Protein Interactions* (MPI’16), Oct. 27–29, 2016, Lawrence, Kansas.
11. “Toward high-throughput interactome-scale protein docking”, *1st International Conference on Computational Genetics and Proteomics*, **invited**, Oct. 18–22, 2016, Guanacaste, Costa Rica.
10. “Modeling protein conformational changes during protein interactions with cNMA — encounter complex-based Normal Mode Analysis”, *6th CAPRI Evaluation Meeting* (CAPRI6), **selected as regular lecture**, Apr. 17–19, 2016, Tel-Aviv, Israel.
9. “Modeling Protein Conformational Change upon Interaction and Mutation with cNMA and MD”, *3rd International Conference on Protein and RNA Structure Prediction*, **invited**, Dec. 14–18, 2015, Punta Cana, Dominican Republic.
8. “cNMA: A Framework of Encounter Complex-based Normal Mode Analysis to Model Conformational Changes in Protein Interactions”, *23rd Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB’15), **selected as proceedings talk**, July 10–15, 2015, Dublin, Ireland.
7. “On conformational changes associated with protein-protein interactions: Normal mode analysis for encounter complexes”, *2014 Protein Folding Conference*, **invited**, July 16–19, 2014, Punta Cana, Dominican Republic.
6. “New results on predicting and disrupting protein-protein interactions”, *Protein Discovery 2013 Summit – Protein-Protein Interaction*, **selected as oral presentation from exemplary submitted abstracts**, Oct. 23–25, 2013, San Diego, California.
5. “Improved flexible refinement of protein docking in CAPRI rounds 22–27”, *5th CAPRI Evaluation Meeting* (CAPRI5), **selected as regular lecture**, Apr. 17–19, 2013, Utrecht, The Netherlands.

4. “Designing and unraveling promiscuous inhibitors against drug-resistant target mutations” (with M Radhakrishnan and B Tidor), *245th ACS National Meeting* (ACS’13-S), **invited**, Apr. 7–11, 2013, New Orleans, Louisiana.
3. “Designing promiscuous inhibitors: Lessons learned from HIV-1 protease inhibition” (with M Radhakrishnan and B Tidor), *241st ACS National Meeting* (ACS’11-S), Mar. 27–31, 2011, Anaheim, California.
2. “Docking refinement by the underestimation of free energy funnels”, **invited**, *4th Conference on Modeling of Protein Interactions* (MPI’07), Sept. 30–Oct. 2, 2007, Lawrence, Kansas.
1. “Structural similarity of binding sites in analogous enzymes” (with D Beglov, R Brenke, D Kozakov, and S Vajda), *234th ACS National Meeting* (ACS’07-F), Aug. 19–23, 2007, Boston, Massachusetts.

• Posters

30. “From Function to Distribution Modeling: A PAC-Generative Approach to Offline Optimization” (with Q Zhang^D, R Zhou, and T Liu), Data-driven and Differentiable Simulations, Surrogates, and Solvers (D3S3) Workshop at the 38th Conference on Neural Information Processing Systems (NeurIPS), Dec. 15, 2024, Vancouver, Canada.
29. “Correlational Lagrangian Schrodinger Bridge: Learning Dynamics with Population-Level Regularization” (with Y You^D and R Zhou), AI for New Drug Modalities (AIDrugX) Workshop at NeurIPS 2024, Dec. 15, 2024, Vancouver, Canada.
28. “Equivariant Blurring Diffusion for Multiscale Generation of Molecular Conformer” (with J Park^P), Machine Learning for Structure Biology (MLSB) Workshop at NeurIPS 2024, Dec. 15, 2024, Vancouver, Canada.
27. “Multi-Modal Contrastive Learning for Proteins by Combining Domain-Informed Views” (with H Xu^M and Y You^D), Machine Learning for Genomics Explorations (MLGenX) Workshop at International Conference on Learning Representations (ICLR), May 11, 2024, Vienna, Austria.
26. “Multimodal Machine Learning for Protein Variant Effect Prediction and De Novo Design” (with M Karimi^D, Y Sun^D, Y Cao^D, and S Zhu^D), *37th Annual Symposium of the Protein Society* (PS’23), Jul. 13–16, 2023, Boston, Massachusetts.
25. “Explainable Prediction of Protein–Ligand Affinity and Generative AI for Protein Design” (with M Karimi^D, D Wu^M, Y You^D, Y Cao^D and S Zhu^D), MIT–MGB AI Cures Conference 2023, Apr. 24, 2023, Cambridge, Massachusetts.
24. “Does Inter-Protein Contact Prediction Benefit from Multi-Modal Data and Auxiliary Tasks?” (with A Talukder^U & Y You^D, the presenters, and others), Machine Learning for Structure Biology (MLSB’22) Workshop at the 36th Conference on Neural Information Processing Systems (NeurIPS), Dec. 3, 2022, New Orleans, USA.
23. “Cross-Modality and Self-Supervised Protein Embedding for Compound–Protein Affinity and Contact Prediction” (with Y You^D), Molecular Machine Learning Conference 2022 (MoML’22), Cot. 21, 2022, MIT, USA.
22. “Noncoding variant effect Prediction using genome sequence and chromatin structure” (with W Tan^D, the presenter), *30th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB’22), July 25–30, 2022, Madison, Wisconsin, USA.
21. “Joint embedding of sequence features (texts) and function labels (graphs) for protein function prediction” (with Y Cao^D), *Learning Meaningful Representation of Life Workshop* (LMRL’21) at NeurIPS 2021, Dec. 14, 2021, virtual.

20. "Embed structure-awareness into protein language models for mutation fitness prediction" (with Y Sun^D, the presenter), *29th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB'21), July 25–30, 2021, virtual.
19. "A hybrid model of genome sequence and chromatin structure for noncoding variants effect prediction" (with W Tan^D, the presenter), *29th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB'21), July 25–30, 2021, virtual.
18. "Cross-modality protein embedding for compound-protein affinity and contact prediction." (with Y You^D, the presenter), *Machine Learning for Structure Biology (MLSB'20) Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS)*, Dec. 12, 2020, virtual.
17. "iCFN: An efficient exact algorithm for multistate protein Design" (with M Karimi), *CAGI* (Critical Assessment of Genome Interpretation) Workshop: Assessing the future of genome interpretation*, Dec. 6-8, 2019, San Francisco, California, USA. **Ph.D. student Karimi was awarded a CAGI* Fellowship (sponsored by an NIH/NHGRI R13).**
16. "Predicting pathogenicity of missense variants with weakly supervised regression" (with Y Cao), *CAGI* (Critical Assessment of Genome Interpretation) Workshop: Assessing the future of genome interpretation*, Dec. 6-8, 2019, San Francisco, California, USA. **Ph.D. student Cao was awarded a CAGI* Fellowship (sponsored by an NIH/NHGRI R13) and selected to give a flash talk.**
15. "De novo protein design of novel folds using guided conditional generative adversarial networks (gcGAN)" (with M Karimi, S Zhu, and Y Cao), *27th Annual International Conference on Intelligent Systems for Molecular Biology* (ISMB'19), July 21–25, 2019, Basel, Switzerland.
14. "Bayesian active learning for optimization and uncertainty quantification in protein docking" (with Y Cao), *9th Conference on Modeling of Protein Interactions* (MPI'18), Nov. 8–10, 2018, Lawrence, Kansas. Also presented at the *CASP (Critical Assessment of Structure Prediction) 13 Meeting*, Dec. 1–4, 2018, Riviera Maya, Mexico.
13. "iCFN: an efficient exact algorithm for multistate protein design" (with M Karimi), *9th Conference on Modeling of Protein Interactions* (MPI'18), Nov. 8–10, 2018, Lawrence, Kansas. Also presented at the *CASP (Critical Assessment of Structure Prediction) 13 Meeting*, Dec. 1–4, 2018, Riviera Maya, Mexico.
12. "On conformational changes associated with protein-protein interactions: Normal mode analysis for encounter complexes" (with T Oliwa), *7th Conference on Modeling of Protein Interactions* (MPI'14), Oct. 23–25, 2014, Lawrence, Kansas. Also presented at the *CASP (Critical Assessment of Structure Prediction) 11 Meeting*, Dec. 7–10, 2014, Riviera Maya, Mexico.
11. "A complex-based normal mode analysis for conformational changes upon protein binding" (with T Oliwa), *10th 3DSIG: ISMB Satellite Meeting on Structural Bioinformatics and Computational Biophysics* (3DSIG'14), Jul. 11–12, 2014, Boston, Massachusetts.
10. "Deciphering and overcoming mutation-based resistance to antiandrogens" (with GL Greene and CL Sawyers et al.), *9th 3DSIG: ISMB Satellite Meeting on Structural Bioinformatics and Computational Biophysics* (3DSIG'13), Jul. 19–20, 2013, Berlin, Germany.
9. "Exploring molecular mechanisms for androgen receptor antagonism and agonism-conferring mutations" (with MD Balbas, MJ Evans, DJ Hosfield, GL Greene, and CL Sawyers), *6th Conference on Modeling of Protein Interactions* (MPI'12), Nov. 8–10, 2012, Lawrence, Kansas.
8. "Targeting drug-resistant HIV-1 protease mutants using inherent protein flexibility" (with B Tidor), *8th 3DSIG: ISMB Satellite Meeting on Structural Bioinformatics and Computational Biophysics* (3DSIG'12), Jul. 13–14, 2012, Long Beach, California.

7. “Designing promiscuous inhibitors: Lessons learned from HIV-1 protease inhibition” (with M Radhakrishnan and B Tidor), *25th Annual Symposium of the Protein Society* (PS’11), Jul. 23–27, 2011, Boston, Massachusetts.
 6. “Designing promiscuous inhibitors: Lessons learned from HIV-1 protease inhibition” (with M Radhakrishnan and B Tidor), *25th Annual Meeting of the Groups Studying the Structures of AIDS-Related Systems and Their Application to Targeted Drug Design*, Mar. 28–30, 2011, National Institutes of Health, Bethesda, Maryland.
 5. “Exploring the role of flexibility in molecular recognition and drug design” (with MK Gilson and B Tidor), *Merck–MIT Computational & Systems Biology Symposium*, Apr. 5, 2010, Boston, Massachusetts.
 4. “Charge optimization of flexible ligands” (with MK Gilson and B Tidor), *23rd Annual Symposium of the Protein Society* (PS’09), Jul. 24–29, 2009, Boston, Massachusetts.
 3. “SDU: Protein docking by the underestimation of binding free energy funnels” (with Ich Paschalidis, P Vakili, and S Vajda), *Thematic Year on Molecular and Cellular Biology, Workshop on Protein Folding*, Institute of Mathematics and Its Applications (IMA’08), Jan. 14–18, 2008, Minneapolis, Minnesota.
 2. “Refinement of rigid-body protein docking predictions using semi-definite underestimation” (with P Vakili, Ich Paschalidis, and S Vajda), **first prize**, *3rd Critical Assessment of PRedicted Interactions Evaluation Meeting* (CAPRI’07), Apr. 20–21, 2007, Toronto, Canada.
 1. “A semi-definite programming based underestimation method for global optimization in molecular docking” (with Ich Paschalidis, P Vakili, and S Vajda), *3rd Conference on Modeling of Protein Interactions in Genomics* (MPI’05), Jun. 26–28, 2005, Lawrence, Kansas.
- Coauthored Abstracts
Presenters are highlighted in bold fonts.
18. **S Irani**, W Tan^D, Q Li, W Toy, C Jones, M Gadiya, A Marra, JA Katzenellenbogen, KE Carlson, BS Katzenellenbogen, Y Shen, and S Chandarlapaty (2024) “An Emerging Class of ER mutations Enhances ER Dimerization and Promotes ER Activity”, *Cancer Research* 84(S9): PO3-29-03. Proceedings of AACR (American Association for Cancer Research) 2024.
 17. **R Taftaf**, X Liu, S Singh, Y Jia, D Scholten, Y Zhang, A Davis, C Reduzzi, Y Cao^D, Y Shen, M Cristofanilli, WA Muller, V Varadan and H Liu (2020) “Single cell RNA sequencing-based identification of molecular drivers in circulating tumor cell cluster formation and lung metastasis”, *Cancer Research* 80(S16) 2603. Proceedings of AACR 2020.
 16. TA Knijnenburg et. al. including Y Shen, M Karimi^D, and H Chen^M (2018), “Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas”, TCGA Legacy: Multi-omic Studies in Cancer (Cell Symposium), Sep. 27–29, Washington, DC, USA.
 15. **W Toy**, KE Carlson, TA Martin, CG Mayne, SW Fanning, P Razavi, J Baselga, Y Shen, GL Greene, B Katzenellenbogen, J Katzenellenbogen, S Chandarlapaty (2017) “ESR1 mutations activate and confer hormone resistance via distinct mechanisms”, *Cancer Research* 77(S13) 999. Proceedings of AACR 2017.
 14. **SW Fanning**, V Dharmarajan, CG Mayne, W Toy, KE Carlson, TA Martin, J Nowak, J Nwachukwu, DJ Hosfield, E Tajkhorshid, S Chandarlapaty, P Griffin, Y Shen, JA Katzenellenbogen, GL Greene (2016) “Bazedoxifene inhibits ESR1 somatic mutants with improved potency compared to tamoxifene and raloxifene”, *Cancer Research* 76(S14) 4854. Proceedings of AACR 2016.

13. **PA Watson**, MD Balbas, Z Zhang, TF Ishmael, KE Lawrence, J Wongivpat, SD Sawyers, Y-M Wu, D Robinson, Y Shen, AM Chinnaiyan, CL Sawyers (2016) “Androgen insensitivity syndrome germline loss-of-function mutations in the androgen receptor that acquire somatic gain-of-function in prostate cancer”, *Cancer Research* 76(S14) 238. Proceedings of AACR 2016.
12. **SW Fanning**, C Mayne, W Toy, Y Shen, A Sharma, S Panchamukhi, J Nowak, KW Nettles, S Chandarlapaty, JA Katzenellenbogen, and GL Greene (2015) “Determining the Role of Somatic ER α Mutations in Acquired Hormone (or SERM) Resistance”, *Cancer Research* 75(S9) P3-05-07. Proceedings of the 37th Annual CTRC (Cancer Therapy Research Center) – AACR (American Association for Cancer Research) San Antonio Breast Cancer Symposium: Dec. 9-13, 2014; San Antonio, TX
11. “A cell-based mutagenesis screen identifies a novel AR mutation that confers agonism to enzalutamide (MDV3100)” (with **MD Balbas**, MJ Evans, DJ Hosfield, GL Greene, and CL Sawyers), *CSHL Meeting on Nuclear Receptors & Disease*, Nov. 2012, Cold Spring Harbor, New York.
10. “Predictions with PIPER, SDU and stability analysis at CAPRI, and our server ClusPro” (with **S Vajda** et al.), *4th CAPRI Evaluation Meeting* (CAPRI4), Dec. 9–11, 2009, Barcelona, Spain.
9. “A multi-step approach to protein–protein docking” (with D Kozakov, R Brenke, D Beglov, P Vakili, ICh Paschalidis, and **S Vajda**), *234th ACS National Meeting* (ACS’07-F), Aug. 19–23, 2007, Boston, Massachusetts.
8. “Integrated understanding of the metabolic and gene regulatory systems of *Shewanella oneidensis* MR-1” (with **QK Beg**, DK Byrne, ME Driscoll, FS Juhn, JJ Faith, ICh Paschalidis, TS Gardner and D Segrè), *108th Annual Meeting of the American Society of Microbiology* (ASM’08), poster, Jun. 1–5, 2008, Boston, Massachusetts.
7. “Systematic identification of regulatory mapping and optimal metabolic engineering strategies in *Shewanella oneidensis* MR-1” (with **DK Byrne**, QK Beg, ME Driscoll, FS Juhn, JJ Faith, ICh Paschalidis, D Segrè, and TS Gardner), *U.S. Department of Energy Joint Genomics: GTL Awardee Workshop VI and Metabolic Engineering 2008* (GTL’08), poster, Feb. 10–13, 2008, Bethesda, Maryland.
6. “Identification of druggable binding sites of proteins by computational fragment mapping using novel FFT correlation based algorithm” (with R Brenke, **D Kozakov**, D Beglov, G-Y Chuang, MR Landon, and S Vajda), *15th Annual International Conference on ISMB & 6th ECCB* (ISMB’07), Jul. 21–25, 2007, Vienna, Austria.
5. “Density-based characterization of hot spots by computational solvent mapping of proteins and applications to accurate three-dimensional pharmacophore generation” (with **MR Landon**, D Kozakov, R Brenke, T Baumeister, F Guarnieri, and S Vajda), *15th Annual International Conference on ISMB & 6th ECCB* (ISMB’07), poster, Jul. 21–25, 2007, Vienna, Austria.
4. “New approaches to protein-protein docking” (with D Kozakov, R Brenke, SR Comeau, ICh Paschalidis, P Vakili, and **S Vajda**), *3DSIG: Structural Bioinformatics & Computational Biophysics Satellite Meeting, 15th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB) & 6th European Conference on Computational Biology (ECCB)* (3DSIG’07), Jul. 19–20, 2007, Vienna, Austria.
3. “Docking by PIPER and refinement by SDU in rounds 6–11 of CAPRI” (with D Kozakov, R Brenke, SR Comeau, D Beglov, and **S Vajda**); and “ClusPro: Performance in CAPRI rounds 6–11 and the new server” (with **SR Comeau**, **R Brenke**, D Kozakov, D Beglov, and S Vajda);

- invited**, *3rd Critical Assessment of PRedicted Interactions (CAPRI) Evaluation Meeting (CAPRI3)*, Apr. 20–21, 2007, Toronto, Canada.
2. “Identification and characterization of hot spot regions in druggable protein binding sites by computational solvent mapping” (with **MR Landon**, S Vajda, D Kozakov, R Brenke, D Lancia, T Baumeister, and S Thiel), *233rd American Chemical Society National Meeting (ACS’07-S)*, Mar. 25–29, 2007, Chicago, Illinois.
 1. “New approaches to protein-protein docking” (with S Vajda, **D Kozakov**, R Brenke, S Comeau, P Vakili, and ICh Paschalidis), *232nd American Chemical Society National Meeting (ACS’06-F)*, Sep. 10–14, 2006, San Francisco, California.

Other Invited Talks

16. “Forward Prediction and Inverse Design of Proteins in Cancer Research”, 7th Annual Symposium Cancer Research: Basic Science to Bioinformatics, Texas A&M University, 10/11/2024
15. “Protein Design Driven by Principles and Data”, Lawrence Berkeley National Laboratory Computational Biosciences Group, 6/13/2023
14. “Algorithmic Protein Design: From Exact Optimization to Generative AI”, Computer Science Departmental Speaker, The University of Texas at Austin, 4/18/2023.
13. “Graph Representation Learning, Generation, and Application in Network Medicine”, Network Science Institute, Northeastern University, 3/31/2023.
12. “Forward Prediction and Inverse Design for Molecular Discovery”, Amazon BioSciences Colloquium (ABC), virtual, 7/28/2022.
11. “AI-Powered Ligand Discovery and Protein Variant Effect Prediction”, Institute of Biosciences and Technology (IBT), Texas A&M University, virtual, 2/10/2022.
10. “Forward Prediction and Inverse Design for Molecular Medicine”, Center for Infectious and Inflammatory Diseases (CIID) Seminar Series, Institute of Biosciences and Technology (IBT), Texas A&M University, virtual, 11/15/2021.
9. “Data-Driven De Novo Protein Design”, Southeast Center for Mathematics and Biology, Georgia Institute of Technology, virtual, 06/25/2021.
8. “Forward Prediction and Inverse Design for Molecular Discovery”, Bioinformatics and Computational Biology Online Seminar Series, Department of Computer Science, Purdue University, virtual, 03/11/2021.
7. “Mechanistic protein modeling for rational drug discovery”, Center for Computational Systems Biology, Prairie View A&M University, Prairie View, Texas, 11/07/2019.
6. “Interpretable Deep Learning for Predicting Compound–Protein Affinity”, Centre for Molecular Informatics, University of Cambridge, Cambridge, UK, 04/08/2019.
5. “Unraveling and Deciphering Resistance Mutations through Modeling and Designing Protein Interactions”, Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, 07/11/2018.

4. “Keeping One Step Ahead of Cancers: Anticipating and Deciphering Resistance Mutations”, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, 12/28/2017; and Institute for Interdisciplinary Information Sciences (IIIS), Tsinghua University, Beijing, China, 12/27/2017.
3. “Inverse Problems in Bioinformatics and Machine Learning”, Institute for Interdisciplinary Information Sciences (IIIS), Tsinghua University, Beijing, China, 12/26/2017.
2. “New Approaches for Prediction and Design of Protein Interactions”, Centre for Integrative Systems Biology and Bioinformatics, Imperial College London, London, UK, 7/8/2015.
1. “Predicting and Designing Protein Interactions Through Optimization”, Key Laboratory of Systems Biology Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China, 8/29/2012.

Dissertations/Theses Supervised

15. Shaowen Zhu*, “Sequence-Structure Codesign for Functional Proteins with Diffusion Models”, Ph.D. in Electrical Engineering, Texas A&M University. * Final exam scheduled early December 2024.
14. Qiang Zhang*, “Machine Learning Approaches to Offline Black-Box Optimization”, Ph.D. in Electrical Engineering, Texas A&M University, December 2024. * Co-advised with Prof. Tie Liu.
13. Yuning You, “Generalizable Graph AI for Biomedicine: Data-Driven Self-Supervision and Principled Regularization”, Ph.D. in Electrical Engineering, Texas A&M University, August 2024.
12. Shreeman Kuppa Jayaram, “Molecular Discovery Chatbot”, B.S. in Computer Engineering, Texas A&M University, May 2024.
11. Yuanfei Sun, “Protein Multimodal Learning for Variant Effect Prediction and Protein Engineering”, Ph.D. in Electrical Engineering, Texas A&M University, December 2023.
10. Wuwei Tan, “Genetic Mutation Effect Prediction: Incorporating 1D Sequence and 3D Structure”, Ph.D. in Electrical Engineering, Texas A&M University, August 2023.
9. Yue Cao, “Optimization, Learning, and Generation for Proteins: Docking Structures and Mapping Sequence–Function Relationships”, Ph.D. in Electrical Engineering, Texas A&M University, December 2021.
8. Maxwell Huffman, “Interconnected Financial Prediction using Time-Series and Network Data”, B.S. in Electrical Engineering, Texas A&M University, December 2021.
7. Arghamitra Talukder, “Multimodal Data Fusion and Machine Learning for Deciphering Protein–Protein Interactions”, B.S. in Electrical Engineering, Texas A&M University, May 2021.
6. Mostafa Karimi, “Circumventing Drug Resistance by Molecular Design: Exact Combinatorial Optimization and Deep Generative Models”, Ph.D. in Electrical Engineering, Texas A&M University, December 2020.
5. Fangtong Zhou, “Identifying Nuclear Receptor Ligands through Sequence-based Deep Learning”, M.S. in Electrical Engineering, Texas A&M University, May 2020.

4. Shaowen Zhu, “De Novo Protein Design of Novel Folds using Guided Conditional Wasserstein Generative Adversarial Networks (gcWGAN)”, M.S. in Electrical Engineering, Texas A&M University, August 2019.
3. Di Wu, “Structured Sparsity Learning for Coevolution-Based Protein Contact Prediction”, M.S. in Computer Engineering, Texas A&M University, December 2018.
2. Haoran Chen, “Improving Molecular-Level Protein Docking and Interpreting System-Level Cancer Mechanism through Machine Learning”, M.S. in Electrical Engineering, Texas A&M University, August 2018.
1. Yi Yan, “Efficiency Prediction and Mechanism Discovery for the CRISPR-Cas9 System”, M.S. in Electrical Engineering, Texas A&M University, May 2018.