

CHRIS L. TANG

gmail @chris.l.tang
New York, NY
917.701.8528

github @[christang](https://github.com/christang)
twitter @[digital_cee](https://twitter.com/digital_cee)
linkedin @[compbio](https://www.linkedin.com/company/compbio)

EDUCATION

Columbia University, New York, NY
Ph.D. Computational Biology (2008), M.Phil. (2003), M.A. (2002)
[UCSF, briefly: 1997]

Massachusetts Institute of Technology, Cambridge, MA
S.B. Biology and Minor in Writing (1997)
Robert A. Boit Prize for Writing; Sigma Xi

Stuyvesant High School, New York, NY
National Honors Society (1993)

EXPERIENCE

Entangible Inc.
Principal Engineer, Founder (2013 – current)

Provided full-stack software engineering services, responsive web application development, and strategic guidance using Django, PostgreSQL, AngularJS, Bootstrap, Celery and Heroku.

Knewton Inc.
Software Engineer, Adaptive Learning (2010 – 2013)

Platform Engineering, DevOps

Primary engineer responsible for Java-based web services for delivering recommendations, and receiving and storing tracking information for recommendations delivered and viewed. Responsible for protocol design of the tracking information, instrumentation, unit and performance testing, configuration, and deployment of these services, as well as maintenance of environment on which web services were deployed.

Maintained tools for providing reports on the calibration quality of the IRT psychometric models used in student math and verbal assessment. Continuously monitored and addressed issues related to the student assessment service for the Knewton test preparation platform. Performed on-call operations support duties as Production Support engineer at least 28 days per year.

Java/Python technology stack included google-guice, Jersey, Cassandra, Hector, ZooKeeper, Apache Thrift serialization, Graphite, kestrel, kafka, rabbitMQ, nginx, Amazon EC2, S3, ELB, DropWizard, CloudFormation, Chef, git. 1-week sprints, unit tests, code review and continuous deployment were supported supported by Jira, Jenkins, Gerrit, and maven. Test preparation platform included MySQL, Ruby on Rails, and Javascript.

Team Lead on Data Science, Analytics

Primary engineer responsible for tracking and analyzing data inputs and recommendation service outputs for validating recommendation quality and correctness based on agreed upon rule-sets and business logic. Mentored team of two analysts responsible for responding to issue reports with Knewton's key business partners and content providers. Managed team of two engineers by performing iteration planning as one of four Team Leads.

Analytics feature developer of student progress on-track/off-track indicator, and student cohort clustering and analysis. Interviewed candidates and collaborated in developing screening criteria for recruiting data science team. Helped specify and contribute to the proficiency and scoring models for the recommendation service. Developed visualization tools for the concept prerequisite graph.

Recognized as a Hack Day winner for contributions made to simulating students working through a knowledge graph.

Tickreel
Machine Learning Specialist (2009 – 2010)

Wrote and trained a C++98/Boost implementation of a Bernoulli n -gram model naïve Bayes classifier for sentiment analysis as part of backend

system for semantic opinion mining. Methods achieved >80% accuracy on published test sets through combination of multi-round training, parameterization and feature set optimization.

Columbia University
Graduate Research Associate (2000 – 2008)

Barry Honig, Ph.D. Director, Center for Computational Biology and Bioinformatics; Professor, Department of Biochemistry and Molecular Biophysics; Investigator, Howard Hughes Medical Institute

Wrote, parameterized and validated statistical models for sequence optimal and suboptimal alignments based on variations on Fischer linear discriminants and generalized log-likelihood ratio methods in multiple settings. Alignments were central to on-going research into detecting faint homology between distant protein families, functional analysis of proteins, and protein structure prediction. Methods were validated and achieved top 5% placement in biennial double-blind prediction experiments in 2002.

Anna Marie Pyle, Ph.D. Professor, Department of Biochemistry and Molecular Biophysics; Investigator, Howard Hughes Medical Institute

Applied and extended the theory of nonlinear Poisson-Boltzmann calculations and Monte Carlo methods to applications where nonlinear effects dominate. Theory and calculations were applied to RNA to support on-going studies to understand a chemical mechanism of enzymatic catalysis that was previously unexpected in nucleic acid. Contributed to computational visualization of electrostatic properties in applications written in Visual C++ and Fortran, leading to the seminal paper describing computational prediction of pK shifts in nucleic acids for the first time.

William Stafford Noble née William Noble Grundy, Ph.D. Assistant Professor, Department of Computer Science

Applied and investigated methods for gene classification using support vector machines and generative models such as position-specific scoring matrices. Implemented and experimented with hidden Markov models and generalized suffix trees for the prediction of secondary structure in proteins.

Millennium Pharmaceuticals, Inc.
Asst. Software Engineer, Process Technology (2000)

Developed computer-control workflows for high-throughput experiments and data storage using Visual Basic, SQL Server, COM and CORBA objects.

Research Associate, Protein Biochemistry (1998 – 2000)
Research Assistant, Protein Biochemistry (Summer 1997)

Primary creator of a process for identifying and quantifying protein differences by writing software to align time-varying gradients in liquid chromatography-mass spectrometry data. Validated methodology by performing controlled experiments in the laboratory.

Designed and developed software to automate the processing of raw mass spectra data, used for analyzing proteins, amino-acid mutations and post-translational modifications.

UCSF, Biophysics Group
Graduate Research Associate (Fall 1997)

Alan Frankel, Ph.D. Professor, Department of Biochemistry and Biophysics

Performed molecular experiments on the structure-based computer design of retroviral RNA.

MIT, Department of Chemistry
Undergraduate Research Student (1995 – 1997)

Jamie Williamson, Ph.D. Associate Professor, Department of Chemistry

Thesis: “Comparison of genetic algorithms to distance geometry for deriving protein models from NMR data.”

Created a computer program for building computer-based structural model of biomolecules using a customized genetic algorithm library in C, and creating UNIX scripts to execute commands in XPLOR and InsightII on an SGI.

Rowland Institute at Harvard
Programmer/Intern (Summer 1994)

Created a computer program for graphically displaying functional MRI images using C and Xaw.

Mount Sinai School of Medicine
Volunteer Research Student (1991 – 1992)

Gillian Small, Ph.D. Professor, Cell Biology & Anatomy

TECHNICAL STACKS

Core languages: Python, Java, JavaScript, Ruby
Analytics: NumPy, SciPy, Matlab/Octave, R
Core frameworks: DropWizard, Jersey, Django, Rails, AngularJS
Core data stores: MySQL, PostgreSQL, Cassandra
Core testing: unittest, junit, rspec, grinder
Core build chain: maven, yeoman, grunt, git, south
Middleware: RabbitMQ, kestrel, kafka, celery
Toolsets: AWS, Ubuntu, Heroku, Sun Grid Engine
K&R, BS: C++98/C, STL, Boost, X11R6/motif, valgrind, template generics
Esoteria: Scala, Scheme, Racket, Perl, CGI, PHP, VB, Fortran

PUBLICATIONS

Tang CL. (2008) Dissertation title: “Computational Studies of Proteins and Nucleic Acids: On pKa Calculations in RNA and the Use of Structure to Improve Sequence Alignments.”

Tang CL, Alexov E, Pyle AM, Honig B. (2007) Calculation of pKas in RNA: On the structural origins and functional roles of protonated nucleotides. *J. Mol. Biol.* 366(5):1475-96.

Forrest L, Tang CL, Honig B. (2006) On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys. J.* 91(2):508-17.

Murray PS, Li Z, Wang J, Tang CL, Honig B, Murray D. (2005) Retroviral matrix domains share electrostatic homology: Models for membrane binding function throughout the viral life cycle. *Structure* 13(10):1521-31.

Tang CL, Xie L, Koh I, Posy S, Alexov E and Honig B. (2003) On the role of structural information in remote homology detection and sequence alignment: New methods using hybrid sequence profiles. *J. Mol. Biol.* 334:1043-1062.

Petrey D, Xiang X, Tang CL, Xie L, Gimpelev M, Mitros T, Soto CS, Goldsmith-Fischman S, Kernytsky A, Schlessinger A, Koh I, Alexov E and Honig B (2003) Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins* 53:430-435.

Pavlidis P, Tang CL, Noble WS (2001) “Classification of genes using probabilistic models of microarray expression profiles.” *Proceedings of BIOKDD*.

Pevzner PA, Mulyukov Z, Dancik V, Tang CL. (2001) Efficiency of database search for identification of mutated and modified proteins via mass spectrometry. *Gen. Res.* 11(2):290-9.

Pevzner PA, Dancik V, Tang CL. (2000) Mutation-tolerant protein identification by mass spectrometry. *J. Comp. Biol.* 7(6):777-87.

INVITED TALKS (*-presenting author)

“From RNAs to pKas: Theory and application to the hepatitis delta virus ribozyme.” *Tang CL**, *Alexov E, Pyle AM, Honig B*. Center for Structural Biology, Yale University, May 2004.

“Sequence alignment and fold recognition using HMAP.” *Tang CL* et al*. SAC-CASP5, December 2002.

SELECTED ABSTRACTS

“Structural determinants of pKa shifts in RNA.” *Tang CL**, *Alexov E, Pyle AM, Honig B*. 13th Annual International Conference on Intelligent Systems for Molecular Biology, 2005.

“Combining alignment sampling and ab initio methods for comparative modeling and fold recognition.” *Petrey DS, Fasnacht M, Forrest L, Kosloff M, Posy S, Tang CL, Pincus D, Li X, Zhu J, Soto CS, Bertonati C, Goldsmith-Fishman S, Friesner R, Honig B*. CASP6, 2004.

“Use of limited suboptimal alignment in homology modeling.” *Tang CL**, *Petrey DS, Fasnacht M, Kosloff M, Alexov E, Honig B*. RECOMB 2004.