



TEMPLE UNIVERSITY
A Commonwealth University

College of Science and Technology
Department of Chemistry

Beury Hall (016-00)
Philadelphia, Pennsylvania 19122-2585
U.S.A.
Phone: (215) 204-7118
Fax: (215) 204-1532

December 9, 2004

Biocomplexity Faculty Search Committee,
c / o Prof. Rob de Ruyter van Steveninck,
Department of Physics,
Indiana University,
Swain Hall West 117,
Bloomington IN, 47405-7105

Dear Prof. Steveninck,

I write in support of Dr. Junwen Wang's application for a faculty position in your department. Dr. Wang earned his Ph.D. in Fishery Science from the University of Washington in 2000. He joined my laboratory in November 2001 after a brief venture at Rawhide Internet Services, where he gained extensive computer programming skills. Dr. Wang has a strong desire to pursue a career in bioinformatics.

Dr. Wang was highly recommended by his graduate thesis advisor, Prof. George Pigott, who until recently had been the director of the Institute for Food Science and Technology at University of Washington in Seattle. His research work had involved in investigating the process of animal protein hydrolyzation. One of the projects focused on processing seafood waste into food seasoning ingredients. In addition to the experimental work on biochemistry of proteins, Dr. Wang also independently carried out computer modeling and simulation of the process of protein hydralyzation.

After joining my lab, Dr. Wang took a special interest in protein structure prediction. He continued the work of protein sequence analysis by using the neighbor-dependent sequence analysis method, a project initiated by a former postdoc in the lab (who has since joined the neural informatic center at Yale University). Unlike the traditional means of determining the preference of amino acids adopting a particular secondary structure in proteins based their frequency of occurrence in the structure database, the neighbor-dependent sequence analysis calculated the effect of neighboring amino acid type on the propensity of residues for adopting a particular secondary structure in proteins. This method produced an enhanced statistical significance scale that allowed us to explore the positional preference of amino acids for protein secondary structures. Dr. Wang showed that the amino acid pair preference for α -helix had its unique patterns and that such patterns were not always predictable by assuming proportional contributions from the propensity values of the individual amino acids. His analysis also yielded a series of sequence patterns that had strong preference for α -helix and β -strand conformations. The details of his work was published in the journal *Protein Engineering*.

The effectiveness of comparative protein structure prediction relies heavily on the sequence alignment between the template and the query sequences. Most of the available algorithms performed poorly when aligning proteins sharing low percentage of sequence identity. Dr. Wang initiated a project aiming to improve the accuracy of aligning remotely related protein sequences. He developed Neighbor-dependent Assisted Sequence Alignment (NdASA). This algorithm incorporated the sequence patterns derived from the neighbor-dependent sequence analysis into a global sequence alignment algorithm. Using the alignment derived from structure-structure alignment as a standard, the NdASA demonstrated superior performance over existing algorithms. For proteins sharing sequence identity of 20 - 25%, the NdASA could improve alignment accuracy over the PSI-Blast by an average of 5%; for more remotely related proteins sharing sequence identity of 13 - 18%, the NdASA could improve the alignment accuracy over the PSI-Blast by an average of 12%. Dr. Wang has developed a web server for protein sequence alignment using the NdASA. The server is now available for public access at <http://astro.temple.edu/~feng/Servers/BioinformaticServers.htm>. A paper describing his approach on NdASA will be published later this year in the journal *Proteins: Structure Function and Bioinformatics*.

Since moving to Upenn, Dr. Wang has continued his aggressively style of tackling difficult biological problems. He has applied the same "neighbor-dependent" principle to computationally identify genomic locations of gene transcription and promoter sites. He has developed novel Markov models that were shown to be effective in identifying gene core promoters. It is my understanding that Dr. Wang has been very productive in the past 10 months at Upenn. I am sure his current mentor, Prof. Hannenhalli, will give you a more detailed evaluation of his current work.

In summary, Dr. Wang has extensive programming skills. Over the course of his training in my lab, he has gained significant knowledge on how to apply computational skills to solve biological problems. In an effort to improve his knowledge base, he also attended some graduate courses offered on campus, including Structure/Function of Proteins and Algorithms in Bioinformatics. Dr. Wang interacts well with people in the group, always willing to assist other graduate students in the lab. Professionally, I think he has the capability to succeed in bioinformatics. I recommend him highly.

Sincerely,



Jan Feng, Ph.D.
Associate Professor of Chemistry
Department of Chemistry
Temple University