My teaching experience began early in college. **During my second year of college I worked as a teaching assistant** with Dr. Oscar Díaz. For three semesters I taught general biology lab to engineering students, which involved preparing practical laboratory sessions demonstrating cellular processes, preparing guides and correcting quizzes. Later in my degree program **I worked as an English teacher for low-income adults** at the AICC Institute in Santiago. Most of my students were working as babysitters or cleaning homes. It was so gratifying to pass on my knowledge in English to those people, knowing that maybe one day they would be considered for higher work positions. In addition, I worked as a Chemistry tutor. These were very rewarding experiences, and they influenced my goals and aspirations a lot.

Aside from teaching, my research experiences also started early. In my third year of college (after a Genetics class with Dr. David Holmes) I joined the Laboratory of Bioinformatics and Genome Biology at the University of Santiago de Chile. The lab was working on genomics of the bacterium *Acidithiobacillus ferrooxidans* involved in acid mine drainage (please see my research proposal for a description of this term). I participated in the annotation project of the bacterium. The annotation was done manually using bidirectional analysis, checking similarity with orthologs and looking for functional domains or motifs. This was my first experience working in research as part of a team.

When I was in my fourth year of college **I earned a fellowship granted by Fundación Andes in Chile**. This is a very prestigious fellowship that allowed me to work in the US during the summer of 2003, at Integrated Genomics Inc. in Chicago, Illinois with Dr. Andrei Osterman. I started working on the annotation and assignment of genes to metabolic pathways. I reconstructed the lipopolysaccharide biosynthesis pathway in Gram negative bacteria by searching for the genes known to be involved in this pathway, and developing a flow diagram that was later incorporated into the ERGO platform (ERGO is the interface where sequence analysis is done). Then, my research project involved the analysis of genomes based on the search for possible missing genes (hypothetical genes that might reveal a function after experimental analysis) in three strains of the bacteria *Buchnera* sp. Finding missing genes is very important because once they are verified they can be assigned to pathways for which steps where unknown. I learned how to use bioinformatic platforms not only to study individual genomes, but also to compare multiple organisms based on their genomic information. My work at IG started as part of the annotation team, and concluded with my own project. I went back to Chile and continued taking classes during that year but I knew that I wanted more lab training and new practical knowledge.

Next summer I flew back to Chicago to participate as a **volunteer in the DNA and Pre-implantation Genetic Diagnosis (PGD) lab at the Reproductive Genetic Institute** with Dr. Svetlana Rechitsky. The RGI provides fertility treatments such as in-vitro fertilization, pre-implantation and prenatal diagnosis. My work included assisting the DNA and PGD lab group doing PCR reactions and DNA extractions from blood and amniotic fluid samples for genotyping and medical diagnostic purposes. I performed all the molecular techniques independently.

In March 2004, coming back from Chicago again, I re-joined the Laboratory of Bioinformatics and Genome Biology at the University of Santiago, and later Millennium Institute in Santiago, Chile with Dr. David Holmes. **My undergraduate thesis research involved studying alternative open reading frames (ORFs) in fully annotated genomes**. Given a gene sequence that is read in the +1 frame (by default) there are 5 other frames in which this same sequence could be translated. We called them alternate ORFs. Examples of alternate ORFs are genes encoded in more than one frame in viral genomes. I was part of a team that took

a large set of annotated genes from the public databases and extracted alternate long ORFs, and **I am a coauthor on the publication of this work**. It has been shown that the most favored alternate frame containing ORFs is frame -1. The known gene plus its ORF in frame -1 are called antisense genes. I used computational genomic analysis involving studies of conserved domains and motifs, applied to the research of alternate ORFs in frame -1 in bacterial and archaeal genomes. I found that antisense genes are widely distributed among organisms. This is very important because in the process of genomic annotation we could miss functions that could be encoded in additional frames. My thesis project was an independent work, which included, among other things, learning and using Perl programming.

For more than a year I worked as an assistant specialist at the Geomicrobiology Lab at the University of California in Berkeley with Dr. Jill Banfield. I started studying microbial communities found in acid mine drainage (AMD). Metagenomic sequencing of environmental AMD samples was done, and near-complete genomes of the most abundant organisms are now available in our lab. I independently worked on the annotation and metabolic reconstruction of the bacteria *Leptospirillum rubarum* (group II) and *Leptospirillum ferrodiazotrophum* (group III). The goal is to know what metabolic pathways the organisms have, and eventually to get a detailed overview of the organisms' capacities to live in such extreme environment. I found, for instance, that *L. rubarum* is the first acidophilic bacterium that has the complete pathway for tolerance to salt and temperature stress. During this year I have been doing a comparison between both species of *Leptospirillum*. This has allowed me to find key metabolic pathways that give different adaptability to each organism. In addition, I do wet lab work. I use fluorescent *in situ* hybridization on environmental samples to study microbial composition, and I culture isolates of different organisms from the AMD environment. **I will report my work in the scientific literature in a first author publication**.

My professional goals have been strongly shaped by my experiences, and my scientific interests are tightly linked to my past research work. I will be an environmental microbial researcher and a university professor, both being equally important to me. The graduate program I chose, as well as the NSF fellowship will greatly help me achieve these goals.

PUBLICATIONS

- Comparative population proteogenomic analyses of coexisting "*Leptospirillum rubarum*" and "*Leptospirillum ferrodiazotrophum*" populations that dominate acid mine drainage biofilms. **Daniela S Goltsman**, Jill F Banfield (**in preparation**)
- Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. Lo I, Denef VJ, Verberkmoes NC, Shah MB, **Goltsman D**, DiBartolo G, Tyson GW, Allen EE, Ram RJ, Detter JC, Richardson P, Thelen MP, Hettich RL, Banfield JF. **Nature**. 2007 Mar 29;446(7135):537-41
- Large-scale, multi-genome analysis of alternate open reading frames in bacteria and archaea. Felipe Veloso, Gonzalo Riadi, **Daniela Aliaga**, Ryan Lieph and David S. Holmes. **OMICS**. 2005 Spring;9(1):91-105.

TALKS

- Biological Analysis of Alternate Open Reading Frames: Fertile Grounds for the Development of New Genes. Aliaga D, Veloso F, Riadi G, Lieph R, Holmes DS. Bioinformatic Workshop, Talca, Chile. Nov. 2004.