Molecular and Biochemical Studies of the Cercarial Proteinase of *Schistosoma mansoni* **MR-19**

James H. McKerrow, Ph.D., M.D.
Department of Pathology, University of California, San Francisco

We have been studying the molecular basis of cercarial invasion of skin. Our hypothesis is that invasion of skin is facilitated by the secretion of a proteolytic enzyme. In previous work, we identified the major secreted protease of schistosome cercariae as a serine protease released from acetabular glands in response to skin lipid. We found that it could degrade macromolecular barriers found in skin such as keratin, laminin, type IV collagen, fibronectin, and elastin. We later purified this enzyme and began to determine the binding specificity of its active site using synthetic peptide substrates and inhibitors. Our more recent work has addressed the following issues:

1. What is the three dimensional structure of this enzyme and can we relate active site structure to the broad substrate specificity and tissue degrading capacity exhibited by the proteinase?
2. How is this enzyme evolutionarily related to other members of the serine protease family like the elastases of higher vertebrates?
3. Can specific inhibitors of this enzyme in fact arrest larval penetration of skin?
4. When is this enzyme synthesized and where is it stored and released?
5. Do other species of schistosomes have a similar enzyme and how much structural variation has occurred during schistosome evolution?

To obtain more structural information about the enzyme, we utilized an amino terminal amino acid sequence determined from the purified proteinase to synthesize an oligonucleotide probe for isolation of a full length cDNA. This work, done in collaboration with Dr. George Newport, currently at the Department of Pharmaceutical Chemistry at UCSF, resulted in the elucidation of the complete amino acid sequence of the enzyme. One of the observations I was able to make was that the cercarial enzyme had regions of sequence identity to other members of the eukaryotic serine protease family, especially pancreatic elastases I and II. This correlated with the ability of the enzyme to degrade elastin. Furthermore, the enzyme appeared to be synthesized as a zymogen like most mammalian serine proteases.

Our most recent work has taken advantage of the deduced amino acid sequence of the protease, and recent advances in computer modeling of the three dimensional structure of proteins developed by one of my collaborators, Dr. Fred Cohen of the Department of Pharmaceutical Chemistry at UCSF. Fred and his graduate student, Lydia Gregoret, produced a three dimensional model of the cercarial proteinase that we are currently utilizing to obtain specific information about the substrate binding specificity of the enzyme's active site. Our experiments with this model are focused around two major goals. The first is to confirm the predictions of the model by site directed mutagenesis of putative beta strands or loop structures. The second goal is to focus on the P1 to P4 binding sites (where the four amino acids on the amino terminal of the peptide bond to be cleaved are bound) to determine the fine structure of substrate binding. These latter experiments utilize both site directed mutagenesis as well as testing of synthetic peptides or inhibitors. We then relate the results back to our primary hypothesis. That is, if we can identify a very specific inhibitor for this enzyme, which has no other observable toxicity to cercariae, we can test whether this inhibitor will block skin invasion.

The other recent advance we have made was in expression of active proteinase in a bacterial expression system using the plasmid pTRAP. This is a system developed by our collaborator, Charles Craik of the Department of Biochemistry at UCSF, to express vertebrate trypsin. It overcomes the difficulty inherent in expressing an active proteolytic enzyme within an organism by directing the product of expression to the periplasmic space of *E. coli* where it will

not damage the vector that is producing it. A postdoc in my laboratory, Johnny Railey, has expressed active enzyme and initiated the first site directed mutagenesis studies suggested by the three dimensional computer model.

By *in situ* hybridization with the cDNA clone for the cercarial protease, Matt Petitt, a technician in my laboratory, identified two cells in the posterior portion of developing cercariae in which messenger RNA for the proteinase was being synthesized. The subsequent differentiation of these cells can be traced to the same acetabular glands from which we originally purified the enzyme. Therefore, we have been able to come full circle from identifying the proteinase in the cercarial secretions, purifying it, developing a molecular probe, and then using that molecular probe to identify the cells from which the glands differentiate. We are following up these studies to also define when the zymogen form of the enzyme is first activated.

Finally, we have begun studies on the molecular evolution of the cercarial protease in different species of schistosomes. Payman Amiri has identified the corresponding proteinase in the non-human schistosome parasite *S. douthitti*. This is one of the causative agents of "swimmer's itch." In contrast to the *S. mansoni* enzyme, the *S. douthitti* enzyme has a markedly different substrate specificity towards peptides with a positively charged amino acid at the P1 site. Recently, using generic molecular probes for serine proteases developed by a postdoc in my laboratory, Judy Sakanari, we have also isolated gene fragments of the cercarial proteinases of *S. japonicum* and *S. douthitti*.