Discovering the determinants of outer membrane protein insertion is a task of basic scientific interest and technological value. It is an investigation with potential for insight both into the molecular mechanisms of bacterial function, and into the chemistry of biological membranes. The technological value comes from the prospect of designing mutants of soluble proteins that will insert into membranes. Outer membrane proteins have a β-barrel fold that is similar to that of many soluble proteins, most importantly GFP. If the determinants of outer membrane protein insertion were fully understood, it may be possible to design a mutant GFP along the same principles that would insert into membranes, which would have important applications as a voltage sensor in neuroscience experiments (Blunck et al., 2005).

Outer membrane proteins (OMPs) are insoluble in water, and if unfolded can spontaneously refold and insert into vesicles (Surrey and Jähnig, 1992). Together this suggests that at equilibrium, OMPs are in membranes, rather than being forced in and held there by some kinetic barrier. It is likely, then, that negative of folding is a necessary condition of insertion.

Under the hypothesis that for a whole OMP is a sum of contributions from solvent-exposed residues on its surface, the problem of estimating for an arbitrary OMP becomes the problem of finding free energies of transfer for individual amino acids. Once this is achieved, the calculation of is as simple as summing the transfer energies. There have been three broad categories of approaches to this problem.

One is experimental. 's for each amino acid have been derived through a mutation study (Moon and Fleming, 2011). Another approach is through simulation. Molecular dynamics simulations have been used to derive 's which are very close to those estimated from experiment (Gumbart and Roux, 2012).

This study is concerned with the development of the knowledge-based approach to the problem. In a knowledge-based approach one desires to know the *energy* of a particular state; and, from a structure database, one knows the *frequency* of that state. Some relation between the two is used to deduce the energies from the frequencies, and the resulting energy function is called a *knowledge-based potential*. This study continues development of the Ez