Date: Sun. 15.02.15

The main goal of this report is to overview the method named *Direct Coupling Analysis* to predict the amino acids pair contacts within single protein chain. Briefly speaking, the method of *Direct Coupling Analysis (DCA)* employs the *Ising* model in statistical physics to construct the interactions among site pairs and use some approximation methods to find out the most probable one. This report will be organized as following. First, we overview the background of *Ising* model and its associating generalized one which is *Pott* model. Second, we present the analogy of *Ising/Pott* models to the interaction system among amino acids pairs. After that, we summarize several approaches to estimate the parameters in *Ising/Pott* models. Finally, a modified version of *DCA* based method will be presented as to figure out the contacts between two proteins in a complex.

1 Ising Model

Let Λ be finite set, and denote $\Sigma_\Lambda \overset{def}{=} \{1, 2, ..., q\}^\Lambda$. In particular, q is set to 2 in *Ising* model. Suppose $A = (a_{ij})_{i,j\in\Lambda}$ is real symmetric matrix, and $\mathbf{h} = (h_i)_{i\in\Lambda}$ is a real vector. The *Hamiltonian* of the model corresponding to those parameters is a mapping $H_{A,\mathbf{h}}\colon \Sigma_\Lambda \to \mathbf{R}$ defined by

$$-H_{A,\mathbf{h}}(\sigma) \overset{def}{=} \sum_{i,j\in\Lambda} a_{ij}\sigma_i\sigma_j + \sum_{i\in\Lambda} h_i\sigma_i$$

and the *Gibbs* measure $G_{A,\mathbf{h}}$ on Σ_Λ is defined by

$$G_{\Lambda,A,\mathbf{h}} \overset{def}{=} \frac{1}{Z_{\Lambda,A,\mathbf{h}}} exp\left[-H_{A,\mathbf{h}}(\sigma)\right]$$