**Pseudo-PSSM and Pseudo-PP**

Inspired by Chou's pseudo amino acid (PseAAC), the pseudo-position specific scoring matrix (PsePSSM) has been a widely used descriptor for proteins and we also get the feature pseudo-physicochemical properties (PsePP) in this paper. For the matrix of a query protein, we firstly normalize it between 0 and 1 using sigmoid function as:

$$f(x) = 1/\left(1 + e^{-x}\right) \tag{1}$$

In this paper, x represents PSSM or PP matrix. There, we take the calculation of PsePSSM as an example and the normalized PSSM matrix can be represented as:

$$PSSM^{'} = \begin{pmatrix} p^{'}_{1,1} & p^{'}_{1,2} & \cdots & p^{'}_{1,20} \\ \vdots & \vdots & \vdots & \vdots \\ p^{'}_{i,1} & p^{'}_{i,2} & \cdots & p^{'}_{i,20} \\ \vdots & \vdots & \vdots & \vdots \\ p^{'}_{L,1} & p^{'}_{L,2} & \cdots & p^{'}_{L,20} \end{pmatrix} \tag{2}$$

Then, the PsePSSM features is calculated in two steps: first, we average each column of the PSSM$^{'}$ matrix, and get a 20-dimensional vector $\overline{PSSM}$ as:

$$\overline{PSSM} = \frac{1}{L}\sum_{i=1}^{L} p^{'}_{i,j}, \quad j = 1, 2, \cdots, 20 \tag{3}$$

Second, we calculate $PSSM^{\phi}$ as:

$$PSSM^{\phi} = \frac{1}{L-\phi}\sum_{i=1}^{L-\phi}\left(PSSM^{'}(i,j) - PSSM^{'}(i+\phi, j)\right)^2, \quad j = 1, 2, \cdots, 20 \tag{4}$$

where $0 < \phi < \Phi$. We set $\Phi$ to 30 in our experiment and get a matrix with dimensions $30 \times 20$.

Finally, the feature $F_{PsePSSM}$ with 620-dimensions is obtained and it can be represented as:

$$F_{PsePSSM} = \left(\overline{PSSM}, PSSM^{\phi}\right) \tag{5}$$

For the PP matrix, we do the same process as PSSM matrix, and get a $30 \times 57 + 57 = 1767$ dimensions feature $F_{PsePP}$ as:

$$F_{PsePP} = \left(\overline{PP}, PP^{\phi}\right) \tag{6}$$