

Significance of triple torsional correlations in proteins

Shiyang Long*, Jianwei Wang† and Pu Tian‡

April 17, 2019

*School of Chemistry, Jilin University

†School of Life Science, Jilin University

‡School of Life Science and School Artificial Intelligence, Jilin University, 2699 Qianjin Street, Changchun, China 130012

Appendix

The derivation steps of equation 8 on the mutual information of A and B was calculated as bellow:

$$\begin{aligned}
MI(AB) &= S(A) + S(B) - S(AB) \\
&\approx - \sum_{i=1}^m w(A_i) * f(A_i) * \ln(f(A_i)) - \sum_{i=1}^m w(B_i) * f(B_i) * \ln(f(B_i)) \\
&\quad + \sum_{i=1}^m \sum_{j=1}^m w(A_i) * w(B_j) * f(A_i, B_j) * \ln(f(A_i, B_j)) \\
&= - \sum_{i=1}^m w(A_i) * (p(A_i)/w(A_i)) * \ln(p(A_i)/w(A_i)) \\
&\quad - \sum_{i=1}^m w(B_i) * (p(B_i)/w(B_i)) * \ln(p(B_i)/L(B_i)) \\
&\quad + \sum_{i=1}^m \sum_{j=1}^m w(A_i) * w(B_j) * \left(\frac{p(A_i, B_j)}{w(A_i) * w(B_j)}\right) * \ln\left(\frac{p(A_i, B_j)}{w(A_i) * w(B_j)}\right) \\
&= - \sum_{i=1}^m p(A_i) * \ln(p(A_i)) + \sum_{i=1}^m p(A_i) * \ln(w(A_i)) - \sum_{j=1}^m p(B_i) * \ln(p(B_i)) \\
&\quad + \sum_{i=1}^m p(B_i) * \ln(w(B_i)) + \sum_{i=1}^m \sum_{j=1}^m p(A_i, B_j) * \ln(p(A_i, B_j)) \\
&\quad - \sum_{i=1}^m \sum_{j=1}^m p(A_i, B_j) * \ln(w(A_i) * w(B_j)) \\
&= - \sum_{i=1}^m p(A_i) * \ln(p(A_i)) - \sum_{i=1}^m p(B_i) * \ln(p(B_i)) \\
&\quad + \sum_{i=1}^m \sum_{i=1}^m p(A_i, B_j) * \ln(p(A_i, B_j)) \\
&= S_p(A) + S_p(B) - S_p(AB)
\end{aligned}$$