MULTIPLE STRUCTURE ALIGNMENT BY OPTIMAL RMSD IMPLIES THAT THE AVERAGE STRUCTURE IS A CONSENSUS

Xueyi Wang and Jack Snoeyink
Department of Computer Science, University of North Carolina at Chapel Hill
Chapel Hill, NC, 27599-3175, USA

Root mean square deviation (RMSD) is often used to measure the difference between structures. We show mathematically that, for multiple structure alignment, the minimum RMSD (weighted at aligned positions or unweighted) for all pairs is the same as the RMSD to the average of the structures. Thus, using RMSD implies that the average is a consensus structure. We use this property to validate and improve algorithms for multiple structure alignment. In particular, we establish the properties of the average structure, and show that an iterative algorithm proposed by Sutcliffe and co-authors can find it efficiently — each iteration takes linear time and the number of iterations is small. We explore the residuals after alignment and assign weights to positions to identify aligned cores of structures. Observing this property also calls into question whether global RMSD is the right way to compare multiple protein structures, and guides the search for more local techniques.