Prediction of protein structure and interaction by physics and informatics

Chaok Seok

Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea

Protein structure prediction problem has challenged theoretical and computational physical scientists since the first protein structure was published in 1958. There have been steady progresses in protein structure prediction since then, but major contributions to the progress came from informatics-based approaches rather than from physics-based approaches. Recently, DeepMind's AlphaFold made a further contribution by introducing deep learning to extract structural information from the large sequence database. Meaningful contribution of physics-based approaches began to be made only in 2012 in the field of structure refinement. However, structural improvements that can be achieved by refinement with current physics-based approaches are very limited due to both energy and sampling problems. To overcome this limitation, we are taking an approach that combines physics and informatics, including deep learning. We take similar approaches to predict interactions of proteins with other proteins or small ligands including short peptides and oligosaccharides. Our goal is to develop protein structure modeling techniques that can provide useful predictions even in the absence of available information, although currently available experimental data would play important roles in developing such techniques. Such modeling methods would be very useful for applications to a wide range of biomedical research and drug discovery.