

2019 ML-MSM Meeting - (October 24-25, 2019)
Poster Abstract Submission Form

PI(s) Grant: Markus J. Buehler and David Kaplan

Institution(s): Massachusetts Institute of Technology (MJB) and Tufts University (DK)

Grant Number: NIH U01 EB014976

Title of Grant: Models to Predict Protein Biomaterial Performance

Abstract Title:

Using Artificial Intelligence to Generate *de novo* Thermally Stable Collagen Sequences

Abstract Authors:

Chi-Hua Yu^{1,2}, Om Prakash Narayan³, Rachael Parker³, David L. Kaplan³ and Markus J. Buehler^{1*}

¹Laboratory for Atomistic and Molecular Mechanics (LAMM), Massachusetts Institute of Technology, Cambridge, USA

²Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan

³Department of Biomedical Engineering, Tufts University, Medford, USA

Abstract

We propose a *de novo* framework to generate new, structurally stable collagen sequences by incorporating artificial intelligence and a self-evolutionary algorithm. Collagen is the most important and abundant structural protein in the human body. Many interesting properties of collagen require chemical and structural stability. Collagen unfolds from a triple helix structure to random coil during a temperature interval. The temperature at the midpoint of this transition, T_m , indicates the stability of a collagen triple helix. Many research efforts have sought to elucidate the relationship between collagen sequences and T_m over the past 20 years. However, finding a robust framework to facilitate the design of a specific collagen sequence to yield a specific T_m remains a challenge. To this end, we developed a platform which adopts a deep learning model that is trained with input sequences collected from literature and experiments to predict T_m and a self-evolutionary module to optimize T_m . The research steps are to synthesize collagen and test its stability as guided by AI. This work has the potential to reduce the time for searching desired collagen sequences with experiments, with applications in medical engineering, bioengineering, pharmaceutical industries and other protein classes.