Using Sequence and Structural Information From GFP-Like Proteins To Design A Fluorescent Version Of G2 Domain Of Nidogen

Robert Langan, Marc Zimmer
Department of Chemistry, Connecticut College

Since the development of green fluorescent protein (GFP) as a biomarker through the late 1990s, leading to the 2008 Nobel Prize in Chemistry, the fluorescent proteins have made a space for themselves center-stage in biotechnological research. Structures for 377 of these GFP-like fluorescent proteins are stored online in the Protein Data Bank (PDB). A bioinformatics analysis of these fluorescent proteins downloaded from the PDB reveals what make these proteins fluorescent. After separating wild-type structures from the engineered variants, a multiple sequence alignment was performed to find highly conserved residues within the wild-types, and to find differences in the engineered variants which give them their engineered properties. This work is in the context of the G2 domain of human Nidogen. The structure of the G2 domain of Nidogen contains an 11 stranded beta-barrel, along with a central alpha helix, similar to GFP with the same beta-barrel and central helix, yet is non-fluorescent. Using the results from the bioinformatics analysis, the goal is to find a minimal set of mutations that will make Nidogen fluorescent, as its structure suggests it should. Ultimately, this elucidates the nature of fluorescent proteins and what is at the core of their fluorescent property.