The Protein Conformation Ontology

Morgan Mitchell [1], Lauren M. Wishnie [1] and Alexander D. Diehl [1]

[1] *Department of Biomedical Informatics, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY, 14203, USA*

Abstract

We have developed the Protein Conformation Ontology (PRC) in order to create a comprehensive ontological representation of protein conformations including secondary, tertiary, and quaternary structures. Our goal is to formally define and describe the conformations of proteins including those that adopt alternate conformations based on changing environmental conditions. The PRC has 89 subclasses of protein secondary structure and includes classes for tertiary and quaternary structures. Because proteins have the potential to take on different conformations based on a wide range of conditions in their environment, we represent protein conformations as subclasses of the Basic Formal Ontology 'disposition' class. An initial set of protein secondary structure conformation classes were created based on the Sequence Ontology (SO) class 'polypeptide secondary structure' class hierarchy as well as information from other sources. These classes are defined as types of conformations that capture their three-dimensional structure rather than as sequence features of polypeptides, as in SO. As a use case, we are using these protein secondary structure conformation classes to model the domain structure of the human voltage-gated sodium channel on the tertiary and quaternary levels. The long-term goals of PRC include describing protein domains in terms of their secondary structure conformation and then describing protein structures according to their domain composition and order, potentially in an automated way based on resources like PFAM and UniProt. These ontological representations of protein structures can then be linked to the Protein Ontology classes for the relevant proteins. By creating these ontological representations of protein structures, we can enable comparison and querying of proteins based on their structural elements, both within and across species. We can also represent the various conformations that proteins and protein complexes may adopt under distinct environmental conditions or due to post-translational modifications. Additionally, we can create ontological representations of protein aggregates linked to pathological conditions, such as neurological diseases, that take into account the alternate conformations of the proteins within them.
https://github.com/Buffalo-Ontology-Group/Protein-Conformation-Ontology

Keywords

protein conformation, Protein Ontology, protein structure, sodium channel

ICBO 2022, Month XX–XX, YYYY, City, Country
EMAIL: mm528@buffalo.edu (A. 1); lmwishni@buffalo.edu (A. 2); addiehl@buffalol.edu (A. 3)
ORCID: 0000-0002-9318-128X (A. 1); 0000-0002-7245-3450 (A. 2); 0000-0001-9990-8331 (A. 3)