

Overview of the Protein Blobulator

The Protein Blobulator is an online resource developed at the Brannigan Lab. The goal of The Protein Blobulator is to annotate and illustrate h and p blobs for any given protein sequence, and visualize the location of disease associated SNPs along the sequence. This webtool approach allows for rapid analysis of ensemble effects.

Analyses available within the Protein Blobulator

A Uniprot ID or any sequence given in one letter notation can be accepted as input. The first chart allows the user to see where there are groups of hydrophobic residues stacked together. Mean hydropathy scores are produced as a bar chart where the x axis is the residue number and the y axis is the mean hydropathy associated with each amino acid residue. Mean hydrophobicity ($\langle H \rangle$) at each residue is again defined as the average Kyte-Doolittle [19] score with a window size of 3 residues, scaled to fit between 0 and 1.

User defined blobs

The defined classification of h blobs is determined by the user via sliders. The threshold for how many residues that must be stacked together to be considered a blob can be adjusted via the “minimum blob size” slider. The mean hydropathy score

that each residue must have to be considered a blob is also defined by the slider “hydropathy cutoff”. All of the possible blobulation outputs that are defined by the sliders are precalculated to allow for rapid adjustment of blob visualizations.

Blobs are outputted and colored from two color schemes that have been developed by Das and Pappu (Pappu 2019).

Disease associated SNPs and their location are produced on all of the blobulation outputs, regardless of coloring scheme. The user is able to click on the DA SNP of choice and a new tab will open its information from ClinVar. Viewing the SNPs on the blobs are especially useful for visualizing in what blobs they are enriched.

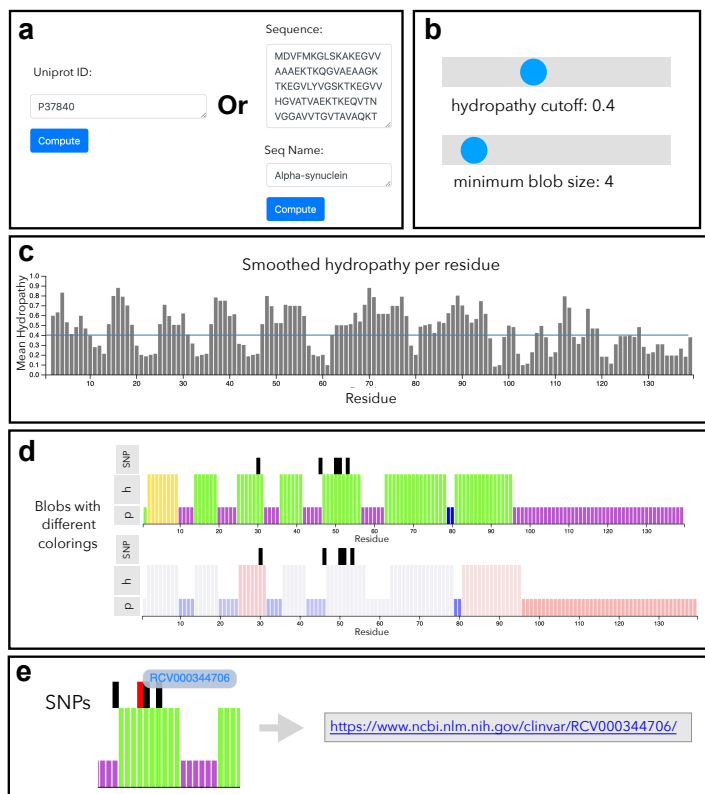


FIGURE xx: Overview of the output generated by the protein blobulator (a) the user inputted data (b) user adjustable sliders (c) linear hydrophobicity plot (d) blobulation output (e) location of DA SNPs along the blobulation output

