Team meeting

Who: modelling Team

Description:

(General) accommodation

* Timescale – want this done before wet lab team go into lab

* If on-board with doing bioinformatics, run this over the summer

* Does Jack have to work solo?

* Schedule separate meetings

* Use DALI search to identify closest match to CH3 domain which contains isopeptide bonding Introduce relevant mutations at those sites

* Or guess somewhere inside the protein? Doubt it will be successful

* Phylogenetic then train machine

* Email rosetta

* Get list of 20 isopeptide-containing proteins from Uli – also ask about electron density, was this from crystallography or can we use pdb

* Then whole genome phylogeny of pili

* Then generate learning set then programme?

Modelling Plan of action:

* Rational Design Use DALI search to align extant isopeptide-containing proteins with the domains of the target antibody. Then cross-compare and see where isopeptide residues are in the extant proteins, then see where we need to introduce Asn/Lys/Glu residues in the target domain

*Do this by automating DALI searching through a Python script, then use the script to screen for ‘Ig’ results. Once screened, manually sift through them. Use this to determine which domain we edit.

Once we identify the residues to edit, hand over to wet lab for site-directed mutagenesis.

This is the first priority, as it gives the wet lab something to do from the start.

* Combinatorial Approach

*dea is to train a neural net to recognise isopeptide-forming sequences. So, we need a training set, and need more than just the 20 extant isopeptide-containing proteins!

Generate the training set using phylogenetic analysis. Study the phylogeny of isopeptide-containing bacteria to predict other proteins in related bacteria which may contain isopeptide bonds (we're assuming the sequences of these are known but isopeptide bonds unproven?). Use this generated set of related sequences and the isopeptide sequences to train the neural net to recognise isopeptide sequences.

Then, use a python script to randomly generate Lys/Asn/Glu mutations in the known sequence of our target domain. Screen this using our neural net for potential isopeptide-containing proteins, then order in 'hit' sequences. Hand over to wet lab for GFP screening.

Time frame for this is tricky as we need to give wet lab time to screen the sequences but also time for Jack to help us with building the neural net. Thus, we suggest commencing the phylogenetics asap.