chargedesign project - theory

renko.devries

March 2023

Hypothesis

For protein sequence blocks used in antifouling applications, we want to minimize non-specific interactions of the designed block with all other types of proteins (say, serum proteins). Key non-specific interaction are the electrostatic interactions. For globular proteins, it is (presumably) required to have many charged residues on the surface to maintain solubility. However, these may easily form small clusters that generate significant local electrostatic surface potentials, and this may lead to non-specific electrostatic interactions with other proteins. We hypothesize that non-specific electrostatic interactions may be minimized by choosing surface residues that generate surface potentials that are as *homogeneous* as possible, as discussed below.

Mean square fluctuations of protein surface potential

At high net charge, non-specific interactions are most likely dominated by the average of the electrostatic potential $\psi(\mathbf{r})$ over the protein surface S:

$$\bar{\psi} = \int_S \psi(\mathbf{r}) dS / \int_S dS \tag{1}$$

Below some absolute value of $\bar{\psi}$, it is not the net charge, but rather the average over the protein surface area S of the fluctuations ("due to charge patches") of the surface electrostatic potential $\Delta\bar{\psi}^2$ that will presumably determine non-specific electrostatic interactions:

$$\Delta\bar{\psi}^2 = \int_S (\psi(\mathbf{r}) - \bar{\psi})^2 dS / \int_S dS \tag{2}$$

We expect that low non-specific interactions always requires that the absolute value of $\bar{\psi}$ is low, hence we will need to focus on finding surface residues that minimize $\Delta\bar{\psi}^2$.

Surface residues that minimize $\Delta\bar{\psi}^2$

As a score function f to be mininized we therefore simply take the fluctuations of the protein surface potential, $\Delta\bar{\psi}^2$,

$$f = \Delta\bar{\psi}^2 \tag{3}$$

which we need to minimize with respect to the identity of the designable surface residues. This we will do using simulated annealing. Suppose that due to a trial move (mutation of designable surface residues) the score function changes by Δf, then the Metropolis acceptance criterion for a simulated annealing move is:

$$\Delta f < 0 \qquad \text{accept} \tag{4}$$

$$\Delta f > 0 \quad \text{accept with probability} \quad \exp(-\Delta f/kT) \tag{5}$$

where kT is a measure for the range of values of the score function f that is sampled. A simulated annealing search would start off at high kT and then slowly move to smaller and smaller values of kT. This need not be done continuously, one can also save the best sequence of a search at high kT, and use that as starting point for a subsequent search with lower kT.