393g Prediction of Pka Shifts in Proteins Using a Discrete Rotamer Search and the Rosetta Energy Function

Ryan Marques Harrison and Jeffrey J. Gray

A pKa value, that is, the pH at which a particular chemical group is equally likely to be protonated or deprotonated, can be shifted relative to standard solvated values due to the chemical group's local environment. For example, a histidine residue surrounded by (usually positively-charged) lysine residues could have its pKa shifted down from the baseline value of 6.3 so that it can maintain a neutral charge and thus avoid repulsive electrostatic interactions with the lysines. pKa values are also used to assess the energy required to place or remove a charge at a particular place on a protein, and thus the prediction of pKa shifts can be a rigorous test of the accuracy of computer models of protein energies. Here, we present a novel method to compute pKa values from the protein structure using methods from the protein structure prediction field. To find the lowest free-energy conformation(s) at a given pH, we use a simulated annealing procedure to search through a rotamer library of typical side-chain conformations including both protonated and deprotonated variants, thus allowing for side-chain relaxation due to charge redistribution. We modified the Rosetta energy function appropriately for the new charged and neutral side chains; the energy function includes van der Waals interactions, an implicit solvation model, hydrogen bonding, generalized Born electrostatics, torsion energies (via rotamer frequencies), and the chemical protonation energy. The algorithm predicts pKa shifts in the third domain of turkey ovomucoid inhibitor and in ribonuclease A to 0.6 and 0.4 mean-squared deviation, respectively, from that measured experimentally over a combined set of 21 aspartic acid, glutamic acid, histidine, and lysine residues. We will also present results for blind predictions on a new, large set of pKa values measured on several mutants of staphylococcal nuclease. This fast and accurate technique opens a range of computational possibilities including docking pH-sensitive proteins, folding proteins under variable pH conditions, modeling pH denaturation, and designing pH sensitive proteins and protein interfaces.