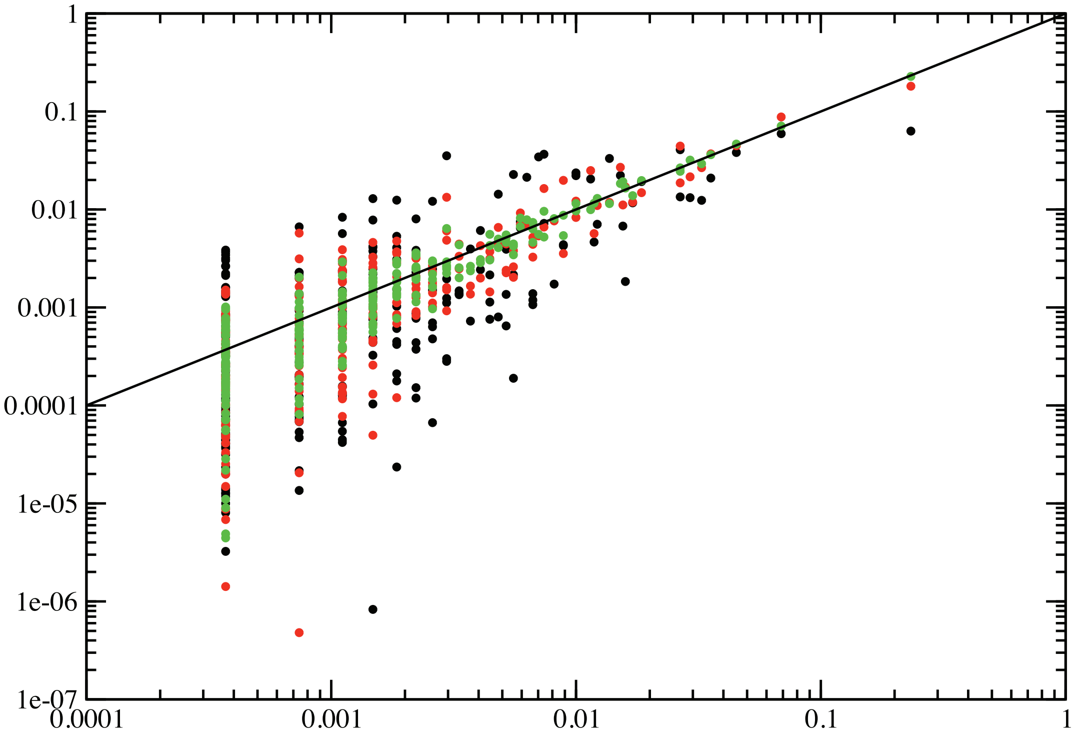

Additional File 2: Predicted vs observed probabilities for the 2^{10} mutational states of the ten-residue group 20–32–46–48–53–54–58–74–82–90 in HIV protease for the PI2+ cohort. Mutational states that were unobserved in the database and would have an observed probability maximum likelihood estimate of zero are not shown. The black dots correspond to the best-fit independent model, the red dots correspond to the best-fit pair-term model (Equation 11), and the green dots correspond to the best-fit three-body model (Equation 12). The solid line of slope 1 corresponds to perfect agreement of the predicted data with the observed, which would be obtained if all higher-order terms were included.