Additional figure 4: χ_1 versus χ_2 histograms for histidine, phenylalanine, glutamate, aspartate and proline. Histograms marked "Training" were generated from the training set; histograms marked "BASILISK" were generated from BASILISK samples.