Amino acids that are involved in binding specificity can be identified with many methods, but few techniques identify the biochemical mechanisms by which they act. We hypothesize that an Analytic Ensemble of techniques, each focused on a single kind of biochemical mechanism that influences binding, could simplify the larger problem into more manageable pieces. Here, we present one technique that can suggest electrostatic mechanisms that influence specificity.

We produced a classifier called DeepVASP-E that applies 3D convolutional neural networks to categorize a voxel-based electrostatic representation of ligand binding sites into categories with different ligand binding preferences. It relies exclusively on voxelized electrostatic data, ensuring that any classification it produces is explained at least in part by electrostatic mechanisms.

We hypothesized that voxels that are salient for classification by DeepVASP-E would also be regions of electrostatic isopotential that are crucial for achieving specific binding. We applied Grad CAM++ for measure classification salience, and then verified the resulting regions against biochemical findings on the proteins in our dataset. Our findings, on two families of proteins with electrostatic influences on specificity, suggest that large salient regions occur nearby identify amino acids that have a substantial electrostatic role in binding. By verifying the explanations generated by our technique against experimentally established explanations in the peer-reviewed literature, we find that our approach can be an effective technique for explaining electrostatic mechanisms that control protein specificity.