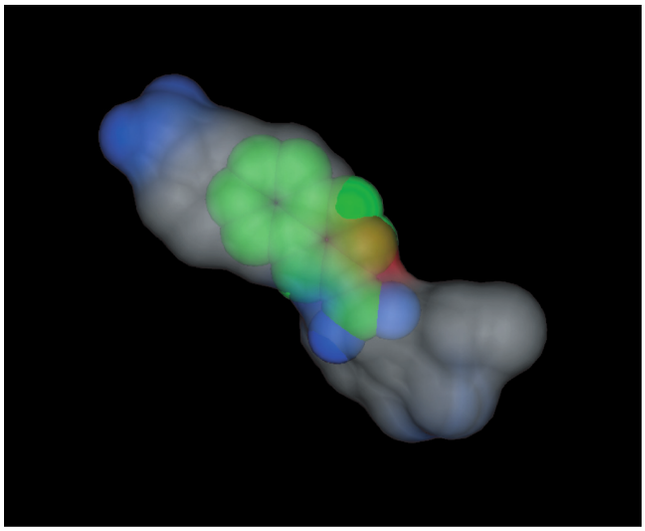

**Figure 7.5.** Rocs overlay to 1o5c ligand does not place small fragment in S1 binding pocket. Transparent surface with gray carbons represents the molecular shape of the 1o5c ligand; solid surface with green carbons represents the molecular shape of compound uk.1-14.

pose had been selected and refined for each ligand in the urokinase and JNK3 data sets. These final docked poses were collated and shown to a local medicinal chemist who has not worked on urokinase or JNK3 program teams. That medicinal chemist visually inspected each docked pose within the context of the binding site environment and offered comments and critiques of the predicted poses; a limited number of docked poses were further refined based on those comments before submitting the predictions to the SAMPL-1 organizers.

### Urokinase plasminogen activator

Urokinase is a serine protease that converts plasminogen to plasmin. I personally have neither worked with nor done computational design for this or any other serine protease, although other members of GSK Computational Chemistry US have supported other serine protease programs. Two public urokinase structures (PDB codes 1o5c and 1owd) were provided as part of the virtual screening challenge and a third structure provided as part of the cross-docking challenge; a sequence- and structure-based overlay of those structures is shown in Figure 7.3. Among the three structures, there was little variation in backbone conformation in the binding-site region [Figure 7.3(a)]; there were small differences in the orientations of some sides chains (not shown), and three residues near the binding pocket were seen in different rotamer states (not shown) – only one of these three was expected to have an appreciable effect on the docking of ligands to the urokinase binding site. Structures 1o5c and 1owd each contained a ligand bound in the protease active site [Figure 7.3(b), 2D structures shown in Figure 7.4]. In both structures the arylamidine arginine mimic binds in the deep S1 pocket, while the bulk of each inhibitor fills the length of the solvent-exposed binding cavity. The amine in the 1owd interacts with an aspartic acid on the protein surface [indicated by a yellow arrow in Figure 7.3(b)]. In both structures, the subpocket marked by a yellow star in Figure 7.3(b) is not filled by any portion of the ligand. When assessing docked poses for any ligands with branched substituents near the arginine mimic, poses that filled this subpocket were manually selected over those that did not fill this region of the binding site.

All of the urokinase ligands to be docked contained some sort of arginine mimic, generally a guanidine or arylamidine although there were two ligands that contained heteroarylamine arginine mimics. Most of the ligands to be docked were large enough to be expected to dock across the entire binding cavity; there were, however, six fragment-sized molecules with molecular weight $\leq$250. The Rocs overlay procedure did not work well for these six molecules, tending to place them in the center of the binding pocket rather than in the S1 pocket (example Rocs overlay shown in Figure 7.5). Docking poses for these six small molecules were generated from a docked pose for a larger molecule containing the relevant S1-binding scaffold; extraneous substituents were removed from the larger docked molecule, and the smaller fragment minimized within the binding site.

For each ligand for which there was a protein/ligand structure, the average rmsd-DPI was computed for all automated predictions; this average is graphed in bold black in Figure 7.6, with 95% confidence intervals shown in dashed black lines; rmsd-DPI values for the manual predictions are shown in bold magenta. The graphed results have been sorted in order of increasing rmsd-DPI for the manual predictions, which has the effect of overemphasizing the manual prediction results; the apparent jaggedness of the automated prediction average is a result of this ordering and has no meaning.

For this serine protease target for which I have no special expertise, the performance for manual and automated predictions is similar. The manual predictions fall inside or

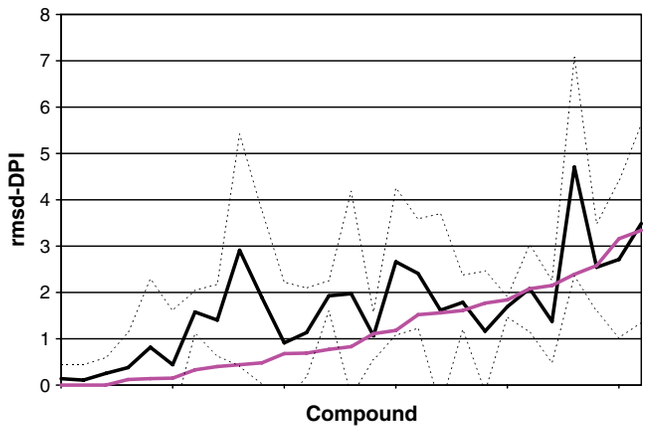


**Figure 7.6.** Results for urokinase docking mode predictions: rmsd-DPI values for manual predictions are shown in bold magenta; mean rmsd-DPI values for automated predictions are shown in bold black with $\pm 1\sigma$ in dashed black lines.