

The Collection is a screening set of 25,000 compounds that has been assembled from the ActiProbe-50 library through Jarvis-Patrick clustering. J-P clustering permits sampling of large library pools through selection of molecules that are representative of a group (cluster) within a library. Molecular characterization is based on 2D fragment descriptors consisting of a central atom and neighboring atoms connected to it within a predefined sphere size -- bonds between central and edge atoms. Groups of molecules are based on the structural similarity of fingerprints (bit strings of these fragments), and the centroid has the highest degree of similarity to other members of a cluster. Chemically unique compounds (singletons) with low similarity to other members of the collection are also identified. Sampling libraries are created from centroid and singleton selection along with augmentation with other members from the largest groups.

Purity of compounds in this, and all our collections, is generally at least 95% as determined by high-field <sup>1</sup>H-NMR and/or MS.

ActiProbe-25K is part of the [ActiProbe Series](#) that includes smaller and larger collections.