

Performance of common similarity measures in virtual screening and lead-hopping

J.C. Baber¹

G.. Tawa², R. Nilakantan³, D. Mobilio³, L. Greenblatt², K. Fan², C. Humblet²

¹ *Chemical & Screening Sciences, Wyeth Research, Cambridge, USA*

² *Chemical & Screening Sciences, Wyeth Research, Princeton, USA*

³ *Chemical & Screening Sciences, Wyeth Research, Pearl River, USA*

A large number of molecular similarity measures are now available to computational chemists and are routinely used in virtual screening exercises – particularly when no structural information is available. Virtual screening may be divided into two general types of problem: the follow up existing hits and lead-hopping to obtain new, structurally distinct, series. In the former compounds that are structurally similar to the query compound are generally sought whereas for lead-hopping compounds that are structurally dissimilar but have similar activity are required.

We will present the result of a recent examination of the virtual screening performance of over 60 different similarity measures in terms of both enrichment and overall performance as defined by the area under a ROC curve. Each method has been tested against the 40 different targets in the Directory of Useful Decoys (DUD) set¹ which provides a diverse range of drug-like classes of compounds. Tests were carried out both exhaustively using each known active in turn as an exemplar and by repeatedly choosing a random training set consisting of approximately 10% of the known actives. In order to test lead-hopping ability the known actives for each of the targets in the DUD set were classified into chemical series by experienced medicinal chemists. The compounds in each series were then used as known exemplars when screening the remainder of the set and the enrichment, overall performance and number of other series identified calculated.

Analysis of these results shows large differences in performance between methods and across targets. As would be expected, increasing the diversity of the test set generally results in a reduction in performance although some similarity methods appear to be more affected by structural diversity than others. This analysis provides a useful benchmark to assess the performance of new similarity methods and may also assist in selecting the most appropriate method, or methods, to use in order to achieve a given set of virtual screening goals.

1. Nuang, N.; Shoichet, B. K.; Irwin, J. J.; Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, 49, 6789-6801