



# Algorithm and Complexity for Designing Multi Route Synthesis

Akavia, A.<sup>1</sup>, Aronowitz, H.<sup>2</sup>, Lerner, A.<sup>1</sup>, Senderowitz, H.<sup>3</sup> and Shamir, R.<sup>1</sup>

<sup>1</sup> School of Computer Science, Tel Aviv University

<sup>2</sup> Intel

<sup>3</sup> Peptor LTD.

It is widely accepted that success of lead identification and optimization is greatly enhanced by synthesizing and screening sets of compounds, which best represent the property space relevant to the biological activity of interest. The two most widely used synthesis schemes are parallel and combinatorial synthesis. Parallel synthesis enables cherry picked sets of compounds to be synthesized. Such sets, if properly selected, adequately represent the property space but are limited in size since each compound is synthesized individually. Combinatorial synthesis enables the production of large sets of molecules in a relatively short time but provides a poorer representation of the property space. The more complex multi route mix & split scheme, introduced by Cohen and Skiena, is a compromise between the parallel and the combinatorial procedures. It combines the advantages of both methods by enabling the rapid synthesis of fairly large sets of compounds while providing a better representation of the property space due to the relaxation of the combinatorial constraint. We propose an algorithm for the design of

complex, multi route mix & split schemes that enable the synthesis of as many of the desired set of compounds as possible, within given resource constraints (*i.e.*, number of colons per step and total number of beads). We generate a graph representation of the synthesis procedure and perform a search for the optimal graph subject to the resource constraints. Each graph is assigned a score based on the number of the desired compounds it produces. We demonstrate the application of the algorithm on sets of compounds selected from a pre-defined property space. This scheme may be combined with diversity or similarity selection methods to provide adequate representation of the property space under given synthesis constraints.

We also explored the complexity of designing multi route mix & split schemes for a given set of desired compounds. Using computational complexity theory tools, we established NP-Completeness and hardness of approximation for several variants of the synthesis design problem.