Preferred Binding Sites of Gene-Regulating Proteins Based on the Deterministic Dead-End Elimination Algorithm

Robert G. Endres

Computational Materials Science Group
Computer Science and Mathematics Division
Oak Ridge National Laboratory
P. O. Box 2008
Oak Ridge, Tennessee 37831

ABSTRACT

Identifying the specific DNA-binding sites of regulatory proteins is essential to understanding gene expression in the biological cell. Unlike statistical bioinformatics tools we choose an atomistic model based on the protein-DNA co-crystal structure and binding energies obtained from the AMBER force field. Our method uses the deterministic dead-end elimination and A* algorithms to search a library of discrete amino-acid sidechain rotamers and DNA base pairs for the optimal binding conformations. As an example, we search all possible 10-basepair DNA sequences (4 to the power of 10) of Zif268, a well-studied mouse regulatory protein.

Research sponsored by the Division of Mathematical Information and Computational Sciences, U. S. Department of Energy. The Oak Ridge National Laboratory is operated by UT-Battelle, LLC for the U. S. Department of Energy under Contract DE-AC05-00OR22725.