CONTEXT SHAPES: A NOVEL APPROACH FOR PARTIAL COMPLEMENTARY SHAPE MATCHING IN PROTEINS

By

Zujun Shentu

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Computer Science

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

Mohammed J. Zaki, Thesis Adviser

Chris Bystroff, Member

Curt M. Breneman, Member

Barb Cutler, Member

Rensselaer Polytechnic Institute
Troy, New York

October 2006
(For Graduation December 2006)

ABSTRACT

Interactions between proteins are governed by the kinetics and stability of the docked conformation, or "pose" . The stability of a pose is governed by the surface electrostatics and by the surface shape. Shapes that are more nearly complimentary will fill space better when docked. The stability of a pose can be approximated by the amount of surface area excluded from solvent, which is approximately the total area over which the two surfaces are complimentary. For this work we assume proteins are static shapes, i.e. rigid-bodies.

To enable interaction-based systems biology studies, we develop a new shape matching approach which successfully docks proteins in a rigid-body sense. When two rigid shapes are positioned as close as possible without passing through each other, there are at least three points of actual contact, called *Physical Contact Points* (*PCP*). If two proteins can be docked, the partial surfaces from the two proteins around any one of the three PCPs should be in high degree of complementarity. We define the notion of *Context Shape* (CS) as the local shape of the protein around a PCP. Superimposing two context shapes implies aligning two surface points. Since surface points are aligned at PCPs, the task of finding PCPs reduces to the problem of evaluating the complementarity of context shapes, which is a rotational alignment problem obviating the need for translational search. We propose to speed up the pose (mapping) computation considerably by pre-computing a *matching table* of all possible mappings. This approach is very efficient because table checking is much faster than rotation operations. Other filters based on solid vectors and solid angles are employed to further reduce the search space.

Our preliminary results indicate that our method finds the crystallographically determined binding interaction within 2.5 Å in a majority of the test cases. Com-

parisons with PatchDock and ZDOCK(PSC) show that our method based on CS gives better predications in both bound-vs-bound and unbound-vs-bound scenarios on a benchmark v2.4 with 84 cases.