

Protein Structure Prediction Using Residual Dipolar Couplings

Sotirios I. Pantos¹ and Ioannis Z. Emiris²

¹*Department of Biology,
National Kapodistrian University of Athens,
Panepistimiopolis 15784, Greece,
e-mail: spantos[at]biol.uoa.gr.*

²*Department of Informatics and Telecommunications,
National Kapodistrian University of Athens,
Panepistimiopolis 15784, Greece,
e-mail: last name[at]di.uoa.gr;*

Abstract. We use efficient algebraic methods and solve the problems of protein structure validation, protein backbone conformation and protein fold by processing Residual Dipolar Coupling (RDC) orientation restraints from solution NMR spectroscopy. RDCs can be recorded and assigned much faster than Nuclear Overhauser Effects (NOEs) required by traditional NMR structure determination methods. Moreover, RDC provides long range restraints suitable for the structure determination of large and extended proteins. In particular, our first application converts RDC equation to a quartic polynomial and obtain discrete solutions. We illustrate our method to ubiquitin (1D3Z) and find that our calculated substructure is in agreement with the NOE structure. Second, we model the protein backbone by drawing upon robot kinematics, and compute the relative orientation of consecutive pairs of peptide planes using RDC orientation constraints; we obtain a unique orientation by considering also NOE distances. Third, we have developed a novel algebraic method for determining the relative orientation of secondary structures, a crucial question in protein fold classification. The orientation of magnetic vector relative to a secondary structure is determined exactly by solving a polynomial system of 3 equations derived from RDC data measured on a single bond vector type (N-H) in two alignment media. Building upon these exact solutions we formulate a polynomial system of 3 equations and find a rotation matrix mapping one secondary structure molecular frame to the other. A unique solution is obtained exclusively from RDC data without the use of short range experimental data such as NOE constraints making our method general. Our algorithms have used robust algebraic operations and have been implemented with the computer algebra system MAPLE.