Extended Abstract
Volume 3:2, 2019

Journal of Genetics and DNA Research

Open Access

Prediction of folding sites of ÃŽÂ-trefoil proteins with irregular structures

Risako Kimura and Takeshi Kikuchi

Ritsumeikan University, Japan., E-mail: sj0029ei@ed.ritsumei.ac.jp

Abstract

Details of protein folding mechanism are still unknown. The solution to this problem is useful for elucidating the mechanism and treatment of diseases caused by misfolding. The relationship between the amino acid sequence and the structure of a protein is generally thought to be higher in structural similarity between proteins with high amino acid sequence identity. However, β-trefoil proteins are known to have a similar structure despite its low sequence identity among super families. In this study, we aim to obtain information on protein folding, targeting β-trefoil protein with a characteristic structure. We already clarified that the central unit is a folding core in β-trefoil protein with high structural symmetry in the previous. In this study, we predicted folding cores for β-trefoil proteins with irregular structures. The compact areas are predicted using a contact map based on inter-residue average distance statistics (average distance map). Then, high interaction residues are predicted by F-value analysis which calculates the contact frequency by using an effective potential derived from inter-residue average distance statistics. From these, we identify the points important in forming the 3D structure along given sequence. We also investigated the conservation of hydrophobic residues among sequences and attempted to clarify residues important for folding of β-trefoil proteins. Furthermore, the folding mechanisms of the β-trefoil proteins are simulated using the Go model and compared it with the obtained results by ADMs. Because of the ADM analyses, compact regions are found in the N-terminal unit and the C-terminal unit in a β-trefoil protein treated in this study. In the result of the F-value analyses, there is a peak of F-value plot in the central unit. After formation of units at both ends folding occurred, suggesting that the central unit interacts with them. Similar results were obtained in the results of the Go model simulations.

Many protein architectures exhibit evidence of internal rotational symmetry postulated to be the result of gene duplication/fusion events involving a primordial polypeptide motif. A common feature of such structures is a domain-swapped arrangement at the interface of the N- and C-termini motifs and postulated to provide cooperative interactions that promote folding and stability. De novo designed symmetric protein architectures have demonstrated an ability to accommodate circular permutation of the N- and C-termini in the overall architecture; however, the folding requirement of the primordial motif is poorly understood, and tolerance to circular permutation is essentially unknown. The β-trefoil protein fold is a threefold-symmetric architecture where the repeating ~42-mer "trefoil-fold" motif assembles via a domain-swapped arrangement. The trefoil-fold structure in isolation exposes considerable hydrophobic area that is otherwise buried in the intact β-trefoil trimeric assembly. The trefoil-fold sequence is not predicted to adopt the trefoil-fold architecture in ab initio folding studies; rather, the predicted fold is closely related to a compact "blade" motif from the β-propeller architecture. Expression of a trefoil-fold sequence and circular permutants shows that only the wild-type N-terminal motif definition yields an intact β-trefoil trimeric assembly, while permutants yield monomers. The results elucidate the folding requirements of the primordial trefoil-fold motif, and also suggest that this motif may sample a compact conformation that limits hydrophobic residue exposure, contains key trefoil-fold structural features, but is more structurally homologous to a β-propeller blade motif.

This work is partly presented at International Conference on STRUCTURAL BIOLOGY AND PROTEOMICS