

Support Vector Machines for predicting protein structural class

ABSTRACT

It is expected that the Support Vector Machine method and the elegant component-coupled method, also named as the covariant discrimination algorithm, if complemented with each other, can provide a powerful computational tool for predicting the structural classes of proteins.

INTRODUCTION

Introduction The observed results by Muskal and Kim suggested that the structural class of a protein might basically depend on its amino acid composition. Many efforts have been made to predict the structural class of a protein based on its amino acid composition. The physical mechanism about this kind of correlation has been discussed by Bahar et al. and Chou. For a systematic description in this area, see a comprehensive review by Chou and Zhang and an updated review. In this paper, we try to apply Vapnik's Support Vector Machine to approach this problem. In this work, Support Vector Machine was performed based on the data sets constructed by Zhou based on SCOP. In ref.19 the reason why these data sets are more reasonable has also been addressed. As a result, high rates of self-consistency and jackknife test were obtained. This has further confirmed that the structural class of a protein is considerably correlated with its amino acid composition.

CONCLUSION

The current study has further supported, from the approach of SVMs, the conclusion drawn by Chou and his co-workers and Zhou that if the coupling effect among different amino acid components can be properly taken into account, the prediction quality of protein structural classes can be significantly improved.