IDENTIFYING TARGET SITES OF DNA-BINDING PROTEINS BY PHYSICAL EFFECTIVE ENERGY FUNCTION. FREE ENERGY ANALYSIS OF λ-REPRESSOR-DNA COMPLEXES.

Tesi presentata da:
Elisabetta Moroni

Tutors:
Prof. Michele Caselle
Prof. Federico Fogolari

Coordinatore del ciclo:
Prof. Federico Bussolino

Anni accademici: 2003 - 2006
Settore scientifico disciplinare di afferenza: FIS/02
Abstract

In all organisms, the expression of specific genes is most commonly regulated at the level of transcription by DNA-binding proteins that also interact with other proteins and often with external signals. For this reason, increasing amounts of data on the structures of DNA-protein complexes are subject of numerous studies to gain insight in the sequence specific recognition event. However the mechanisms underlying the specific binding are still poorly understood.

Several studies showed that there is no clear one-to-one correspondence between amino acids and bases, even if some frequently occurring interaction has been observed, and that their interactions are widely distributed in space, since the same pair may interact using a variety of geometries; vice-versa, the same structural motif can be used by proteins to recognize different DNA-sequences. An important contribution to interaction derives from contact made with the sugar-phosphate backbone of DNA, whereas only one third of all contacts are specific. Moreover these contacts are often mediated by water. Finally binding-induced conformational changes both in protein and in DNA often play some role in the protein-DNA recognition. Inspection of these structures reveals that, although general rules for recognition can be derived, the ambiguous and complex nature of the recognition mechanisms precludes a simple recognition code, therefore the prediction of DNA target sequences is not straightforward.

DNA protein interactions can be studied using different computational methods, which can complement the current experimental methods and offer some advantages, since experimental methods are generally more laborious and slow.

In the present work, physical effective potentials are used to calculate the free energy of binding of the λ repressor-DNA complex, for which structural and thermodynamic experimental data are available.

Using statistical thermodynamics, it is possible to derive the expression of the free energy of binding of two molecules as the sum of the intermolecular energy (evaluated using a molecular mechanics force field), a solvation free energy term and an entropic term. Different solvation models are used including distance dependent dielectric constants, solvent accessible surface tension models and the Generalized Born model. The effect of Molecular
Abstract

Dynamics simulations on the computed binding energies is assessed. After having assessed the accuracy of different methods and protocols by comparison with experimental data, the free energy of binding for non-specific complexes using the best energetic model has been estimated, obtaining good agreement with earlier theoretical suggestions. Moreover, as a results of these analyses, we propose a protocol for the prediction of DNA-binding target sequences. The possibility of searching regulatory elements within the bacteriophage-λ genome using this protocol is explored. Our analysis shows good prediction capabilities, even in absence of any thermodynamic data and information on the naturally recognized sequence. Our results support the conclusion that physics-based methods can offer a useful complement to sequence-based methods, improving their reliability.