

Profili 3D: analisi di superficie proteiche per l'identificazione di determinanti funzionali

A. Via, F. Ferre', B. Brannetti e M. Helmer Citterich

Centro di Bioinformatica Molecolare - Dipartimento di Biologia - Universita' di Roma "Tor Vergata" -
via della Ricerca Scientifica - 00133 Roma

Abbiamo sviluppato una procedura (de Rinaldis et al., 1998) per calcolatore che consente di analizzare e confrontare superficie proteiche che siano associate a specifiche funzioni (es.: particolari abilita' di legame o di catalisi).

Sulla base di un allineamento multiplo di strutture proteiche omofunzionali ed in analogia con il metodo dei profili per la ricerca di omologie di sequenza (Gribskov et al., 1987), il programma permette di realizzare un profilo 3D che puo' essere usato per fare ricerche nel PDB e selezionare proteine con particolari caratteristiche di superficie e funzionali. L'analisi del profilo 3D consente l'individuazione di residui conservati sulla superficie delle proteine sovrapposte e la definizione di determinanti strutturali associati a particolari funzioni biologiche. Questo metodo puo' inoltre essere utilizzato: i) per identificare strutture proteiche che, pur avendo fold differenti, abbiano una o piu' regioni di superficie con proprieta' chimiche e funzionali simili (evoluzione convergente); ii) come sistema esperto per la mutagenesi sito-specifica o iii) per il protein design.

Stiamo utilizzando i profili 3D per analizzare i determinanti di superficie legati a sei motivi del database PROSITE: aminoacyl-transfer RNA sintetasi di classe II, perossidasi, dominio EF-hand di legame al calcio, sito di legame all'eme della famiglia citocromo c, sito di legame a nucleotidi del p loop, sito attivo delle aspartil proteasi eucariotiche e virali. Tali motivi di sequenza non sono in grado di riconoscere tutte e sole le sequenze proteiche associate alla loro funzione, ma selezionano anche falsi positivi e talvolta non selezionano tutte le sequenze cui e' associata la funzione loro assegnata. Nostro scopo e' indagare se invece esista e possa essere identificato un motivo di superficie specifico per ognuna delle funzioni analizzate. Nel caso del p loop, l'analisi dei risultati ha permesso considerazioni di particolare interesse biologico.

L'applicazione di questo metodo all'intero database di strutture note potrebbe consentire la realizzazione di una banca dati di motivi di superficie capaci di identificare tutte e sole le strutture associate ad una specifica funzione, in quei casi in cui l'analisi della sola sequenza non lo consentirebbe. Lo sviluppo dei progetti di "structural genomics" rende molto piu' vasto l'insieme delle proteine cui applicare il metodo.

Gribskov, M., Mclachlan, A.D. and Eisenberg, D. (1987). Profile analysis: detection of distantly related proteins. Proc. Natl. Acad. Sci USA. 84, 4355-4358.

De Rinaldis, M., Ausiello, G., Cesareni, G. and Helmer Citterich, M. (1998) Three-dimensional profiles: a new tool to identify protein surface similarities. J.Mol.Biol., 284,1211-1221.