

Predizione della specificità di riconoscimento dei domini SH3

Barbara Brannetti, Allegra Via, Gianni Cesareni e Manuela Helmer Citterich

Dipartimento di Biologia, Università di Roma "Tor Vergata", tel. 06/72594314

e-mail: barbara@obelix.bio.uniroma2.it

Scopo di questo lavoro è lo sviluppo di un programma per calcolatore in grado di predire la specificità di legame di un dominio SH3 sulla base della sola sequenza. Il database di dati cristallografici comprende una decina di complessi SH3/peptide o proteina. Considerando anche dati sull'interazione tra SH3 e peptidi disponibili in letteratura o ricavati da esperimenti di phage display, possiamo costruire una matrice di frequenze di interazioni residuo-residuo tra posizioni nella sequenza degli SH3 e posizioni nella sequenza del ligando. La matrice dei contatti SH3/peptide può essere utilizzata allo scopo di valutare coppie di sequenze SH3/peptide e predire la loro capacità di legame.

La matrice delle frequenze dei contatti contiene una quantità di informazione ovviamente limitata dal database di interazioni su cui è costruita, inoltre non contiene punteggi negativi. Abbiamo costruito una seconda matrice SH3 specifica, ottenuta dalla matrice delle frequenze dei contatti arricchita con il metodo dei profili (Gribskov et al., 1987). La matrice dei profili è completamente piena (contiene uno score anche per residui non ancora inclusi nel database SH3/peptidi di partenza) e contiene sia valori positivi che negativi.

Entrambe le matrici possono venire usate per produrre funzioni di score e valutare probabilità di legame tra sequenze di domini SH3 e peptidi, sequenze di proteine o intere banche dati di proteine in tempi molto contenuti (circa 30 minuti cpu su una Silicon Graphics R5000 per analizzare l'intera banca dati delle sequenze delle proteine del PDB, circa 14000 sequenze).

Il metodo è stato messo a punto nello studio dell'interazione tra SH3 e peptidi, ma può essere facilmente applicato anche allo studio di diversi domini di interazione protein-proteina (domini SH2, PH, MHC etc.). Stiamo al momento valutando le prestazioni del metodo con le due diverse matrici descritte. E' importante sottolineare che, all'aumentare del database che costituisce la base informativa iniziale (con altre strutture e/o con altri dati di interazione provenienti da letteratura o phage display) ci aspettiamo un consistente aumento del potere predittivo del metodo.

Referenze:

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