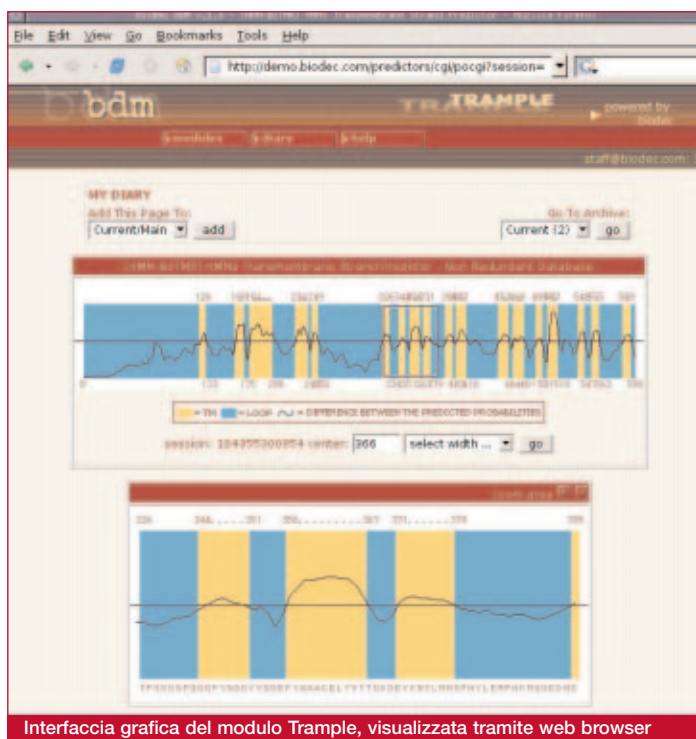


Marco Parenti, Elena Fioravanzo:  
S-IN Soluzioni informatiche Ivan Rossi, Michele Finelli: BioDec

## Dalla sequenza alla struttura in pochi click

Recenti sviluppi tecnologici rendono più semplice l'acquisizione di informazioni strutturali e funzionali fondamentali su target biomolecolari, favorendo il processo di drug discovery



Interfaccia grafica del modulo Trample, visualizzata tramite web browser

Grazie alle attuali tecniche di modellistica molecolare, divenute strumenti di quotidiana applicazione nel settore della ricerca farmaceutica, si è fatta pressante la necessità di disporre di accurate strutture tridimensionali di possibili target biologici. I successi della genomica e della proteomica hanno generato un'enorme quantità di dati sia a livello di codice genetico sia del proteoma in esso codificato.

Si conoscono ormai milioni di sequenze, ma solo di poche migliaia di proteine e' nota la struttura tridimensionale. Il problema di conoscere la sequenza, ma

contenuti i risultati generati dall'applicazione di strumenti di analisi di sequenza proprietari o di altra origine. Sono stati sviluppati diversi moduli: Trample (Trans-Membrane Protein Labelling Environment) basato su artificial neural network e hidden markov model, specifico per l'identificazione e la caratterizzazione di regioni trans-membrana; Angler, classificatore concepito per l'analisi massiva di interi proteomi batterici, ideale per l'identificazione delle proteine associate alla membrana esterna; Tangram, studiato per l'identificazione di omologhi remoti di sequenze e fold

non la funzione o la struttura di una proteina, può essere affrontato mediante l'analisi di sequenze proteiche, che permette di derivare dalla sequenza il massimo delle informazioni possibili.

**BioDec** sviluppa sia strumenti integrati (BioDecoding Machines) per la decodifica e l'annotazione di sequenze proteiche, che database di sequenze annotate dove sono

recognition; Zendock, pensato per la predizione dei siti di interazione proteina-proteina e basato, ancora una volta, su artificial neural network in grado di analizzare la superficie esposta di una struttura proteica ed evidenziare le regioni sulla superficie dove è massima la probabilità di interazione.

Applicando Zendock a tutte le strutture note del Protein Data Bank, è stato generato il database Zenpatches. Dove possibile, i prodotti sono basati su application server, utilizzabili attraverso un'interfaccia web, che ne rende facile e immediato l'utilizzo anche a un utente non specializzato.

### Un software per il Comparative Protein Modelling

I metodi di Comparative Protein Modeling, impiegati per generare strutture tridimensionali di proteine di cui è nota solo la sequenza primaria, si basano sull'osservazione che un'alta similarità di sequenza si riflette generalmente in un'elevata similarità strutturale.

È quindi possibile ipotizzare strutture 3D di proteine in base a strutture note di una o più proteine omologhe. Innovativo, in questo campo, è Prime (Schrodinger, Inc., distribuito in Italia da S-IN Soluzioni Informatiche), un software che combina tecniche di Homology Modeling/Comparative Modeling, utilizzabili solo quando si ha

