

ПРОТОКОЛ

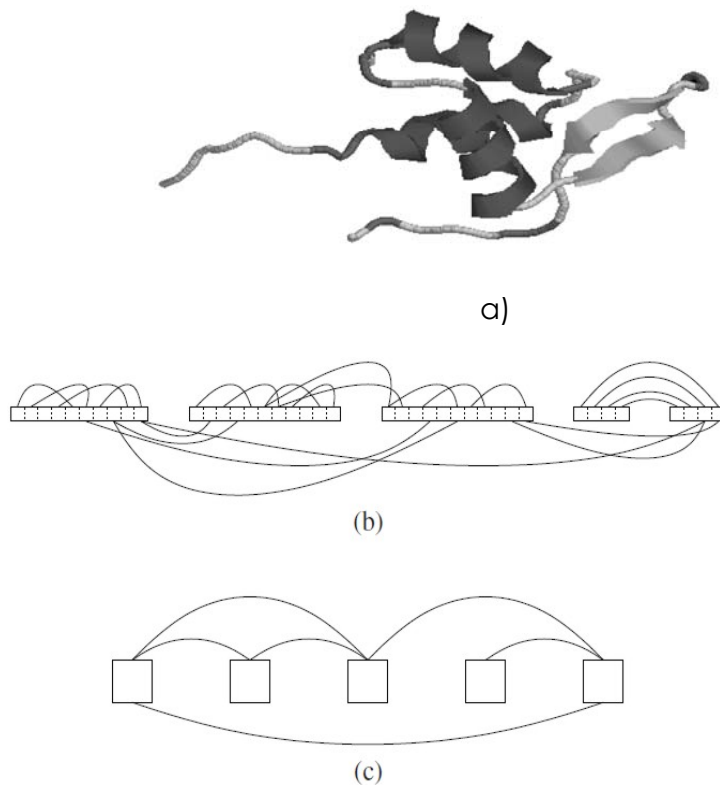
От работната среща на втора работна група по проект ДВУ 01/197 „Биоинформатични изследвания върху структурата и активността на протеините”, проведена в ИМИ – БАН на 20.03.2009 г.

В работната среща участваха проф. Янев, гл. ас Тренчев и гл. ас. Стоян Везенков.

Работната среща беше открита от проф. Янев. Той представи накратко проблематиката, в която е работил: предсказване на 3D Протеинови структури.

Протеиновите структури може да бъдат изследване чрез експериментално методи, като например, или ядрено магнитен резонанс, спектроскопия и др. Въпреки последните постижения в тези области, те са все още скъпо струващи и бавни и тази технология не може да се справи с големият брой белтъци, генерирани в последното десетилетие. Ето защо в молекулярната биология все повече се използват изчислителни модели за определяне на 3D структурата на протеините. Предсказването на 3D структурата на протеините може да бъде формулирана по следния начин: като се има предвид, линейната последователност, която е низ от 20 аминокиселини, да се определят координатите на всяка аминокиселина в 3D форма на протеина. Този проблем в литературата е известен като Protein threading problem (РТР) (вероятното пространствено подреждане на аминокиселините в белтъците).

Проф. Янев представи, както свои разработки по проблематиката, а така също направи обзор на направеното в научните среди като цяло по текущия проблем.



Фиг. 1. А) 3D структура на протерин с β нагъване и цикли б) Съответствия между изследвани участъци в свързан граф с) Оптимизирани съответствия между изследвани участъци в свързан граф

Разгледани бяха няколко математически модела за решаване на РТР проблема. Набелязани бяха конкретни стъпки за обучение на участниците в работната група и навлизането им в научната проблематика.

Проф. Янев предостави научни статии и книги на участниците в работната група за подготовка, с които да се запознаят и останалите членове на работни групи две и три.

Разгледани бяха няколко подходящи софтуера за решаване на оптимизационни задачи.

Обсъдена беше възможността да се включат студенти към работната група. Беше отчетена възможността някои студенти да направят дипломни работи по тематиката на работната група като например разработване на програма за предсказване на 3D структурата на протеини с алгоритъма Flexible Alignments for Protein Threading.

Обсъдени бяха възможностите за работа с биологични бази от данни. Те са големи библиотеки, които съдържат информация от различни научни области, като например: протеомиката, геномика, молекулярна биология и др. Информацията, включена в биологични бази данни, съдържат генетичната локализация (и в клетъчни и хромозомни структури), функция на гена, структура, клинични ефекти на мутации, както и сходствата на биологични секвенции и структури.

Разгледани бяха възможности за работа различни биологични бази от данни. Например

Доклад по темата представи гл. ас Тренчев.

Генетични бази от данни

Най-известните генетична база данни е GenBank, на NCBI (Национален център по биоинформационни технологии). В нея може да бъдат намерени различни нуклеинови и аминокиселини последователности. Възможно е съвместно терсене от NCBI в EMBL (Европейска молекулярна биологична лаборатория), а също и в базата данни DDBJ (ДНК данни банка на Япония).

Протеинови бази от данни

В тези бази от данни са представени нуклеотидни последователности и връзките им с протеинови структури. Представени са характеристика на протеините като 3D структура, хидрофилност, поляритет и др. Най-големи бази от данни са Swiss-Prot, Uni-Prot, и ExPASy.

Представени бяха opensource бази от данни: *BioSQL*, PostgreSQL и MySQL, и езици от високо ниво BioJava, BioPerl и BioPiton..

Решения:

Взето беше решение, че много подходящ за решаване на част от задачите ще бъде оптимизационния пакет CPLEX на фирмата ILOG и да се предложи на академичният съвет на проекта го закупи.

Под ръководството на ръководителя на работна група IV – гл.ас. Иван Тренчев бе взето решение: студенти да направят обзор на съществуващите биологични бази от данни и подходящите за използването в проекта OpenSoure бази от данни.