

## Структурна біоінформатика в постгеномну еру

К. О. Одинець, С. С. Івахно, Д. Б. Ковальський, Б. Т. Токовенко, О. І. Корнелюк

Інститут молекулярної біології і генетики НАН України  
Вул. Академіка Заболотного, 150, Київ, 03143, Україна

---

*Структурна біоінформатика — це нова галузь молекулярної біології, яка використовує методи комп'ютерного аналізу для моделювання просторової структури білків і макромолекулярних комплексів для з'ясування їхніх функцій. Метою структурної біоінформатики є також створення селективних модуляторів функціональної активності білків як нових лікарських препаратів. Прогрес структурної біоінформатики обумовлений бурхливим розвитком геноміки — розшифрування декількох сотень геномів про- та еукаріотних організмів і переходом біології в постгеномну еру. В огляді розглянуто такі питання, як моделювання просторової структури білків, протеоміка та інтерактоміка, біоінформатика при аналізі транскриптому, молекулярна динаміка білків, моделювання взаємодій з лігандами, комп'ютерний дизайн лікарських препаратів тощо. Описано застосування цих підходів при дослідженні тирозил-тРНК синтетази ссавців та протеази вірусу імунodefіциту людини. Подано інформацію про перший український біоінформаційний web-портал BioUA, який створено для проведення досліджень у галузі геноміки та структурної біоінформатики.*

---

Вступ. Біоінформатика — це новітня галузь біології, яка використовує комп'ютерні технології при аналізі та систематизації генетичної інформації для з'ясування структури і функцій макромолекул та складних надмолекулярних комплексів [1—4]. Біоінформатика органічно виокремилася з комп'ютерної біології, де комп'ютери та суперкомп'ютери стали самостійним інструментом пізнання в сучасній молекулярній біології. Важливим чинником бурхливого розвитку біоінформатики стало практичне завершення геномного етапу досліджень в останні 15 років та перехід біології в постгеномну еру. В результаті розшифровки структури геномів вірусів, мікроорганізмів, еукаріотів і геному людини з'явилися гігантські обсяги інформації щодо послідовностей нуклеїнових кислот і білків. Ця інформація стала основою для досліджень наступних рівнів її реалізації, таких як регуляція транскрипції і альтернативний сплайсинг генів, регуляція генних кластерів і генних мереж, трансляція у поліпептидні послідовності та транспорт білків у клітині, метаболічні і регуляторні шляхи клітини, білково-білкові і білково-нуклеїнові взаємодії та ін. Як певний завершальний результат у цьому на-

прямку постає завдання створення комп'ютерних моделей прокаріотної і еукаріотної клітин [5—6]. Для його розв'язання і використовується системна біологія — нова інтегральна біокомп'ютерна галузь, в якій структурна біоінформатика займає одну з найважливіших позицій (рис. 1).

Структурна біоінформатика використовує методи комп'ютерного аналізу для моделювання просторової структури білків і складних макромолекулярних комплексів (білково-білкових, білково-нуклеїнових) та аналізу механізмів молекулярного впізнавання. Метою структурної біоінформатики є також створення селективних модуляторів функціональної активності біополімерів як нових лікарських препаратів (drug design).

В огляді розглянуто окремі аспекти структурної біоінформатики, та наведено приклади досліджень, які проводяться останніми роками у відділі білкової інженерії Інституту молекулярної біології і генетики НАН України.

Геноміка. Першим повністю розшифрованим геномом став геном бактеріофага  $\phi$ X174, який складається всього з 5386 пар нуклеотидів. Значний прогрес у розшифровці геномів почався в 90-х роках минулого століття завдяки розробкам автоматичних секвенаторів нуклеїнових кислот та необхідного програмного забезпечення і створення

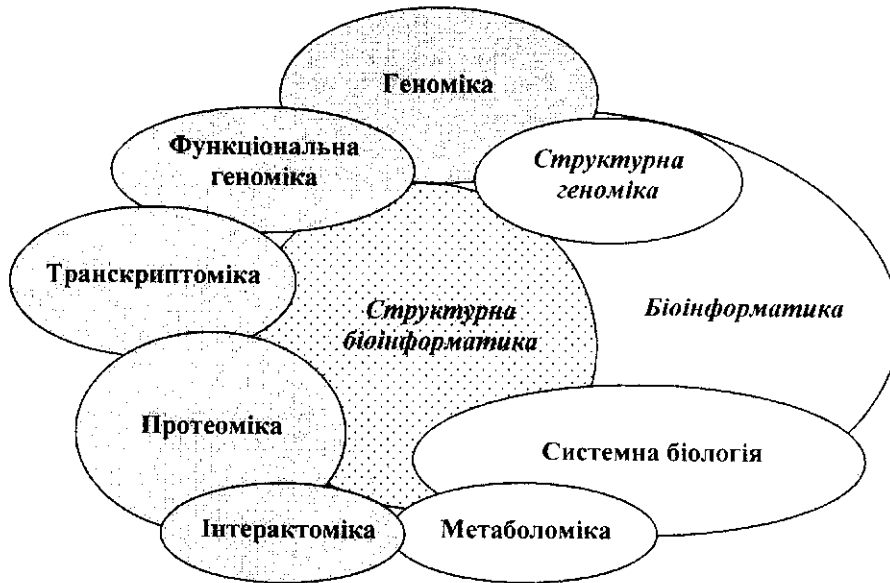


Рис. 1. Місце структурної біоінформатики серед інших галузей системної біології

спеціалізованих центрів секвенування, таких як Sanger Genome Center [7]. Першими організмами, геноми яких було секвеновано, стали бактерія *Haemophilus influenzae* та мікоплазма *Mycoplasma genitalium*, які є патогенами і викликають інфекційні захворювання у людини. Отже, геноміка орієнтована на прикладне застосування геномної інформації і робить значний внесок у вирішення ключових проблем фармакології і медицини.

За переліком бази даних GOLD (Genomes On-Line Database) [8] на березень 2004 р. вже повністю секвеновано 182 геноми (з них 23 еукаріотичних, 18 архебактеріальних і 141 прокаріотичний). На різних етапах завершення знаходяться ще 488 прокаріотичних та 417 еукаріотичних геномів, тобто розшифровка геномів іде наростаючими темпами. Визначено будову геномів пекарських дріжджів *Saccharomyces cerevisiae*, багатоклітинних організмів: нематоди, дрозофіли, арабідопсису і т. д. Найвизначнішим досягненням стала розшифровка геному людини у 2001 році. Картографування і аналіз усіх генів дозволяють порівнювати їх з генами інших організмів, визначати зміни в процесі еволюції та при певних захворюваннях.

Геномна послідовність містить план будови організму, однак її реалізація потребує функціонування різноманітних білків. Одною з основних особливостей геномних даних є значна кількість генів з невідомою функцією або з низькою гомологією до послідовностей з інших організмів. Для деяких геномів ця величина перевищує 40 % [9]. Крім того, виявлено велику кількість ізологічних генів. Відповідно перед біологією постає масштабна

задача ідентифікації та визначення клітинних функцій продуктів цих невідомих генів. У першу чергу це стосується генів, пов'язаних із захворюваннями, оскільки специфічні функції значної кількості їхніх білкових продуктів все ще залишаються невідомими.

Анотація геномів включає визначення функцій генних продуктів і проводиться на основі переважно їхніх амінокислотних послідовностей. У тих випадках, коли амінокислотної послідовності білків недостатньо, встановлення просторової структури білків допомагає передбачити їхні функції. Цей напрямок покладено в основу структурної геноміки, метою якої є отримання структурної інформації для нових, не охарактеризованих раніше білків. Ідентифікація білків, яка базується на порівнянні просторових структур гомологічних білків, часто є більш успішною. Це пояснюється тим, що в багатьох випадках еволюція консервувала подібність просторових згорток набагато довше, ніж гомологію амінокислотних послідовностей.

Передбачення функцій білків виходячи з їхніх послідовностей і просторових структур залишається складною проблемою, оскільки навіть гомологічні білки часто виконують різні функції і, крім того, значна кількість білків виконує ще й додаткові (неканонічні) функції. Методи ідентифікації гомології послідовностей і/або структур між білком з невідомою функцією і добре охарактеризованими білками або виявлення мотивів, консервативних для певних родин білків, все ще далекі від однозначного апіорного встановлення функцій.

Іншими методами біоінформатики, які дозво-

ляють отримати додаткову інформацію, є виявлення і прогнозування паттернів білково-білкових взаємодій, пошук віддалених гомологів, передбачення типу просторової згортки та кореляції між наявністю гомологічних білків у різних організмів з відомими функціональними властивостями. Також проводиться ідентифікація структурних і функціональних мотивів та моделювання просторової структури окремих доменів, білків і макромолекулярних комплексів.

Обсяг накопиченої інформації про геномні послідовності значно перевищує можливості дослідників з їхнього аналізу та використання, тому необхідною стала розробка нових методів зберігання геномної інформації та методів комп'ютерного аналізу геномів.

Моделювання просторової структури білків. Визначення просторової структури білків є необхідним етапом для встановлення взаємозв'язку між їхньою структурою і функціями. Як правило, природна еволюційно відібрана амінокислотна (АК) послідовність за фізіологічних умов набуває досить визначеної просторової структури (точніше, набір конформацій), яка детермінує біологічну функцію даного білка. На березень 2004 р. відомо вже 24,9 тис. експериментально визначених структур [10]. З іншого боку, успіхи у секвенуванні геномів обумовили постійно зростаючий потік даних про АК-послідовності сотень тисяч різноманітних білків і лише для невеликої кількості з них (менше 5 %) доступні експериментальні структурні дані. Тому все більш значущими стають комп'ютерні методи передбачення просторової структури білків, які належать до методів структурної біології *in silico*.

Сучасні потужні комп'ютерні програми можуть враховувати квантово-механічні ефекти у великих молекулярних системах там, де це є необхідним (розрив та утворення ковалентних зв'язків тощо), але для моделювання структури білків у більшості програм можливості обмежуються емпіричними методами молекулярної механіки і вони не враховують електронної будови атомів.

На основі статистичного аналізу тисяч експериментально визначених структур білків було розроблено загальні емпіричні правила і алгоритми, які покладені в основу сучасних методів передбачення і моделювання їхньої просторової структури. Ці підходи розвинулися як допоміжні при уточненні структур, одержаних методами рентгеноструктурного аналізу (РСА) та спектроскопії ядерно-магнітного резонансу (ЯМР), що дало змогу встановити загальні правила пакування поліпептидних ланцюгів та методи їхнього конформацій-

ного аналізу, оптимізації просторової геометрії і мінімізації вільної енергії.

Для передбачення просторової конформації білків доцільно використовувати інформацію про білки, для яких вже встановлено тривимірну структуру. Для побудови структурної моделі білка застосовують, як правило, найефективніший метод моделювання за гомологією, який ґрунтується на фундаментальному принципі взаємозв'язку між рівнем гомології АК-послідовностей білків та подібністю їхньої просторової структури. Чим вищий рівень гомології, тим подібніші тривимірні конформації, тобто просторова структура білків краще зберігається протягом еволюції, ніж АК-послідовність, оскільки для природного добору є важливішою просторова структура молекули білка, яка визначає його функцію і завдяки цьому є більш консервативною. Крім того, білки можуть виявляти значну схожість просторових структур без явної гомології АК-послідовностей. Це можна пояснити існуванням обмеженої кількості різних типів згорток білків (яка оцінюється приблизно в 2—3 тисячі), тому навіть негомологічні білки можуть мати ідентичний тип просторової згортки (можлива конвергентна еволюція).

Якщо для білка, структуру якого моделюють (білок-мішень), існує гомологічний білок, для якого вже експериментально встановлено просторову структуру, то атомні координати останнього можна використати як просторову матрицю для моделювання за гомологією. Цей підхід включає, як правило, такі етапи: 1) пошук у банках даних гомологічних білків-шаблонів (матриць) з експериментально визначеною просторовою структурою; 2) вирівнювання АК-послідовності білка-мішені з послідовностями одного чи кількох білків-матриць (якщо гомологія послідовностей становить більше 40 % ідентичності, то якість вирівнювання послідовностей визначає якість моделі); 3) корекція вирівнювання, особливо на границях делецій чи вставок; 4) генерація ковалентно неперервного ланцюга (каркаса) моделі на основі вирівнювання (це може бути каркас найгомологічнішої матриці або гібридний каркас з декількох матриць); 5) генерація «канонічних» поверхневих петель (у випадку наявності інсерцій або делецій) шляхом отримання з відповідних банків даних або добудова петель методами *ab initio* (петлі в основному визначають конформаційні відмінності гомологічних білків); 6) «вбудовування» бічних радикалів у каркас та оптимізація їхньої геометрії (для консервативних амінокислотних залишків, як правило, зберігають початкові конформації); 7) мінімізація вільної енергії всієї моделі, іноді з використанням

молекулярної динаміки у водному оточенні; 8) перевірка моделі вибіркоким повторенням попередніх етапів.

Чим вищий рівень гомології з білком-матрицею, тим краща якість отриманої моделі просторової структури. Модель структури білка, побудована на основі матриці з гомологією більше 90 %, частіше за все, дуже подібна за якістю до кристалографічних структур. Якщо рівень гомології з білком-матрицею становить 75 % і більше, іноді потрібно додавати вставки поверхневих петель, але вони часто є дуже короткими. Якщо матриця має 50 % ідентичності з послідовністю білка-мішені, то середньоквадратичне відхилення (RMSD) між структурами майже завжди становить приблизно 1,5 Å (із значно більшими локальними похибками). Геометрія структури попередньо отриманої моделі білка (промоделі) має бути оптимізована, тобто потрібно знайти локальний чи глобальний мінімум вільної енергії даної структури. Мінімізація енергії є оціночною процедурою, призначеною для пошуку найімовірнішої конформації молекули з мінімумом вільної енергії. Подібну процедуру покладено в основу уточнення кристалографічних структур білків, визначених методом PCA, а також моделювання фолдингу їхніх макромолекул.

Сервери і web-сайти із структурної біоінформатики. В галузі біоінформатики налічується сотні web-сайтів і серверів. Серед основних світових центрів біоінформатики слід вказати на такі: NCBI [11], EBI і ExPASy [12]. Інформація щодо визначених послідовностей (первинних структур) біополімерів зберігається у спеціалізованих банках даних, для генів і геномів це банки даних GenBank [13], EMBL [14] і DDBJ, для білків — GenPept, TrEMBL [15], Swiss-Prot [15] та ін. Дані стосовно просторових структур білків, визначених методами PCA і ЯМР, депонуються і зберігаються у банках даних просторових структур — PDB (Protein Data Bank) [10], MMDV [16] та ін. На середину березня 2004 р. PDB налічує 24,7 тис. експериментально встановлених структур білків, нуклеїнових кислот і макромолекулярних комплексів.

Робота з просторовими структурами неможлива без спеціалізованих програм, серед яких найдоступнішими є Swiss-PDB Viewer [17], DS ViewerPro, MolMol [18], Protein Explorer [19], RasMol [20] та ін. У попередньому огляді авторів [21] наведено дані про інші програми і сервери, які використовують для роботи з просторовими структурами білків. Значна кількість програм присвячена різним видам аналізу АК-послідовностей білків, які можуть дати корисну інформацію для структурного аналізу. Прикладом можуть бути програми

для передбачення вторинної структури, пошуку трансмембранних ділянок, сайтів протеолізу, антигенних детермінант, сигнальних пептидів у секреторних білках, специфічних мотивів посттрансляційних модифікацій. Окремі бази даних і програми застосовують для пошуку гомологічних і структуроподібних білків та їхньої систематизації за типами просторових згорток: SCOP [22], CATH [23], CE [24], DALI [25], HSSP [26] тощо.

Моделювання за гомологією можна виконувати безпосередньо в Internet з використанням модельовальних web-серверів, таких як Swiss-Model [27], 3D-JIGSAW [28], Geno3D [29], FAMS [30], Modeller [31], What If [32], SCWRL [33] та ін. За допомогою кожного з цих серверів можна провести пошук білків-матриць, власне процес моделювання, оптимізацію геометрії і перевірку якості отриманих моделей. Серед комерційних пакетів програм слід виокремити Insight II, SYBYL і QUANTA, але значна вартість робить їх практично недосяжними для пересічного користувача.

За відсутності відомих структур з виявленою раніше амінокислотою гомологією здійснюють пошук інших структуроподібних білків та ідентифікацію типу просторової згортки. Ці методи і програми, які використовують різні алгоритми, поєднуються під спільною назвою «threading», тобто методи «протягування нитки». До таких програм можна віднести Eval123D, 3D-PSSM [34], BIO-INBGU, FASS [35], FUGUE [36], LOOPP [37], SAM-T99/T98 і THREADER. Для невеликих білків (до 100 амінокислотних залишків) розпочато створення програм з прогнозування структури *ab initio* (HMMSTR/ROSETTA [38]), але якість отриманих моделей все ще значно поступається методам моделювання за гомологією.

Якість одержаних моделей перевіряють і оцінюють з використанням спеціальних програм, які порівнюють моделі з узагальненими статистичними даними, отриманими з експериментально встановлених структур білків. До таких програм належать: ANOLEA [39], ERRAT [40], PROCHECK [41], Prosa II, PROVE, Verify3D [42] і WhatCheck [43]. Як правило, аналіз здійснюють за допомогою кількох програм, що оцінюють різні характеристики просторової структури білків.

Порівняння структур білків на тривимірному рівні. Створення алгоритмів і програм для просторового порівняння структур білкових молекул дозволило поглибити знання з молекулярної еволюції білків [44]. Алгоритми порівняльного аналізу мають два чітко виражених етапи: 1) використання певної функції для побудови матриць еквівалентних атомів двох білкових молекул; 2) пошук

найоптимальніших субматриць, що дають найвищий коефіцієнт гомології для порівнюваних білків.

Показником гомології двох білкових молекул є параметр середньоквадратичного відхилення в положенні еквівалентних атомів (RMSD, як правило, застосовується для  $C_{\alpha}$ -атомів поліпептидних остовів), хоча при пошуку в базах даних більшість програм наводять показник Z-score, який обчислюється шляхом побудови кривої нормального розподілу RMSD усіх досліджуваних білків. Якщо Z-score = 5, то це свідчить про високий рівень гомології двох білків [45].

Більшість сучасних програм базується на інтрамолекулярному підході для структурного вирівнювання, оскільки він дозволяє проводити швидкий пошук у базах даних з невеликою втратою чутливості (можливість знаходження найвіддаленіших гомологів) [45]. Серед алгоритмів, що застосовують інтрамолекулярний підхід, є оптимізація методом Монте-Карло (програма DALI [25]), комбінаторне продовження (програма CE [24]) та динамічне програмування (програма SSAP [46, 47]). Однією з проблем структурного вирівнювання є порівняння віддалено споріднених білків, які мають велику кількість делецій/інсерцій окремих амінокислотних залишків (звичайно в петлях на поверхні білкової глобули), а також окремих елементів вторинної структури. Для вирішення цієї проблеми застосовують декілька стратегій:

— вилучення з вирівнювання ділянок з нерегулярною вторинною структурою, таких як петлі (метод використовується програмою VAST [48]);

— розділення білкової молекули на фрагменти та порівняння індивідуальних фрагментів з наступним їхнім об'єднанням у повне структурне вирівнювання алгоритмом оптимізації Монте-Карло (гексапептидні фрагменти використовуються програмою DALI [25]);

— врахування делецій/інсерцій шляхом попереднього вирівнювання послідовностей з вилученням структурної інформації методом динамічного програмування (алгоритм запропоновано Расселом і Бартом для програми STAMP [49]).

Покращання швидкості і чутливості алгоритмів спричинило появу численних баз даних, що зберігають класифікацію білкових молекул за доменами та мотивами, які ґрунтуються на структурному вирівнюванні білків, найвідомішими з них є вже згадані SCOP [22] і CATH [23].

Моделювання просторових структур каталітичного та цитокін-подібного модулів цитоплазматичної тирозил-тРНК синтетази ссавців. Прикладами проведення моделювання просторової структури білків за гомологією є виконані нами

роботи з побудови моделей двох основних модулів тирозил-тРНК синтетази (TyrRS) ссавців (бик, *Bos taurus*) [50—52]. Субодиниця цієї TyrRS (528 амінокислотних залишків) складається з  $\text{NH}_2$ -кінцевого каталітичного (N) та  $\text{COOH}$ -кінцевого (C) модулів. N-модуль TyrRS вміщує власне каталітичний домен з топологією згортки Россмана та антикодон-впізнавальний  $\alpha$ -спіральний домен, яких достатньо для ензиматичної активності «мінімальної» форми ферменту в умовах *in vitro* [53—66]. Моделювання здійснювали з використанням програм HyperChem 6.01 [67] (силове поле MM+, алгоритм Pollack-Ribiere) і Modeller 4.0 [31]. В отриманій моделі (рис. 2, а, див. вклейку) консервативні амінокислотні залишки у згортці Россмана (Gly41, His49, Tyr166, Pro167, Gln170, Asp173, Gly185 і Gln188) залучені до формування активного центра синтетази. В  $\alpha$ -спіральному домені високі ступінь консервативності мають залишки Ala248, Cys250, Asn258, Lys282, Gly285, Lys310, а також Lys244, з якого починається лізин-багатий кластер  $^{241}\text{VKKKLLKK}$  цього домену. Цей кластер формує структурну западину, яка є ділянкою можливого зв'язування гомологічної тРНК<sup>Tyr</sup>.

C-модуль цитоплазматичних TyrRS гомологічний цитокіну ЕМАР II ссавців (52,7 % ідентичності) і у вільному стані має схожі біологічні властивості в експериментах *in vitro* [65]. З використанням методів порівняльного моделювання за гомологією побудовано модель просторової структури C-модуля TyrRS бика (рис. 2, б, див. вклейку). Основним структурним елементом C-модуля є домен з типом просторової згортки ОВ-фолд (В-домен, залишки Val363-Lys470), який вміщує п'ять  $\beta$ -складок, характерних для цього типу згортки. Подібно до ряду інших ОВ-фолд-доменів тієї ж надродини, C-модуль має спорідненість до нуклеїнових кислот. Порівняння моделі C-модуля з антикодонзв'язуючими доменами аспартил- і лізил-тРНК синтетаз та рядом інших ОВ-фолд-доменів дозволило передбачити поверхню взаємодії з РНК і функціонально важливі амінокислотні залишки. Найконсервативніший залишок серину (Ser436) знаходиться в структурному мотиві, який, імовірно, специфічно взаємодіє з просторовою структурою тРНК. Структурна модель C-модуля TyrRS є основою для аналізу подвійної функції цього домену: структуроспецифічної взаємодії з тРНК, а також його цитокін-подібної функції.

Аналіз функцій та еволюції білка за просторовою гомологією: А-субдомен цитоплазматичної тирозил-тРНК синтетази ссавців та цитокін ЕМАР II. Цитокін ЕМАР II, який є високогомологічним до C-модуля TyrRS, протеолітично

відщеплюється від попередника ЕМАР II — білка р43 — з утворенням зрілого поліпептиду. ЕМАР II складається з двох доменів — ОВ-фолду (амінокислотні залишки 148—252) та А-субдомену (залишки 253—312) [69]. Як РНК-зв'язуючий домен ОВ-фолд зустрічається у великій кількості прокариотичних та еукариотичних білків, тоді як еволюційне походження А-субдомену досі не встановлено. Окрім TugRS та ЕМАР II, цей домен зустрічається в рослинній метіоніл-тРНК синтетазі та дріжджовому білку Arg1p, який є кофактором MetRS та GluRS [61]. А-субдомен приєднується до ОВ-фолду через лінкер, який обвиває NH<sub>2</sub>-кінець ЕМАР II і включає два витки 3<sub>10</sub>-спіралі. Сам А-домен складається з шести β-структур, зібраних в антипаралельний β-шар (β<sub>c</sub>-β<sub>a</sub>-β<sub>b</sub>-β<sub>d</sub>) та β-шпильки між структурами β<sub>c</sub> і β<sub>d</sub>.

Для вивчення еволюційної консервативності ОВ-фолду і А-субдомену проведено структурне вирівнювання ОВ-фолдів і А-субдоменів ЕМАР II (ідентифікаційний код PDB файлу 1EUI) [69] та TugRS (код 1NTG) [70]. Вирівнювання виявило майже повну ідентичність ОВ-фолдів двох білків (RMSD = 0,6 Å, Z-score = 5,7), у той час як А-субдомени відрізнялися набагато більше (RMSD = 1,3 Å, Z-score = 5,2), маючи дві делеції/інсерції у петельних ділянках. Множинне вирівнювання з використанням 18 послідовностей білків з анотованими А-субдоменами показало, що найвища консервативність амінокислотних залишків припадає на мотив РККК (залишки 265—269 у proЕМАР II), в якому лише білок OSJNB з *Oryza sativa* має змінену послідовність.

На відміну від множинного вирівнювання на рівні послідовностей, порівняння просторової структури білків дозволяє знаходити більш віддалені гомологи. Відомо, що ОВ-фолд, топологічно однаковий з NH<sub>2</sub>-кінцем ЕМАР II, зустрічається у тРНК-зв'язуючих бактеріальних білків Ttbp (*Escherichia coli*) і CsaA (*Thermus thermophilus*) [71]. Однак вирівнювання первинної послідовності (субодиниця А, залишки 1—109) та ЕМАР II показало, що лише 25 % залишків є ідентичними. Структурне порівняння ЕМАР II-домену (код 1EUI) та CsaA (код 1GD7), основане на алгоритмі комбінаторного продовження, виявило високу просторову гомологію між цими білками — RMSD = 1,7 Å; Z-score = 6,0. Подібна топологія спостерігалася не лише для ОВ-фолдів, але й для перших 24 амінокислотних залишків А-субдомену.

Ці залишки формують лінкер між ОВ-фолдом та А-субдоменом ЕМАР II, а у білка CsaA вони беруть участь у димеризації через водневі зв'язки та гідрофобні взаємодії. Оскільки у ЕМАР II-доме-

ну лінкер обвиває NH<sub>2</sub>-кінець ОВ-фолду і знаходиться на поверхні білка (дві 3<sub>10</sub> спіралі і β-структура), можна припустити, що А-субдомен з'явився в процесі злиття двох модулів, один з яких вже знаходився у попередника ЕМАР II. Дані структурного вирівнювання також підтверджуються тестом на консервативність амінокислотних залишків (сервер ConSurf [72]), який виявив консервативні амінокислотні залишки Ile90, Glu91, Pro96, Ser99, Gly102, Ile105 в лінкерній ділянці ЕМАР II-домену.

Біоінформатика в аналізі транскриптому. Одним із центральних напрямків функціональної геноміки є аналіз транскриптому — сукупності молекул РНК, які транскрибуються з геному конкретного типу клітин у певний період диференціації та під впливом факторів довкілля [73]. Розрізняють два підходи до аналізу транскриптому:

1) гібридизація зворотно транскрибованих молекул мРНК, які виступають у ролі проб до чітко охарактеризованого набору кДНК клонів чи олигонуклеотидів, кожен з яких репрезентує окремий транскрипт;

2) визначення кількості транскриптів шляхом безпосереднього секвенування кДНК бібліотек або спеціально виготовлених популяцій мРНК (обчислення кількості SAGE тегів [74] і EST транскриптів).

Перший підхід ґрунтується на гібридизації з використанням ДНК-мікроарей [75]. Залежно від методу нанесення ДНК (контактний чи неконтактний прінтинг) ДНК-мікроарей може містити до 40000 елементів з діаметром кожної плями 100—150 мкм [76—78]. Після гібридизації флуоресцентно мічених молекул мРНК з мікроареем інформація про рівень експресії генів визначається детекцією інтенсивності флуоресценції з кожної плями мікроарей-сканерами [79]. Мікроарей широко застосовуються в різних галузях біології для вивчення рівня транскрипції генів та змін у транскриптомі при патологічних процесах, наприклад, трансформації нормальних клітин у ракові [80]; пошуку генів, специфічних для окремих стадій розвитку та диференціації [81, 82]. Мікроарей-експеримент надає одночасно інформацію про експресію десятків тисяч генів. Для подальшого використання ці дані мають бути трансформовані і статистично проаналізовані. Біоінформатика відіграє важливу роль на всіх етапах мікроарей-експерименту — від дизайну нуклеотидних послідовностей для нанесення на мікроарей до розробки реляційних і об'єктно орієнтованих баз даних для збереження отриманої інформації та методів їхнього аналізу [83].

Детально мікроарей-біоінформатика та її за-

стосування проаналізовано в нашому попередньому огляді [84].

**Протеоміка та інтерактоміка.** Протеом клітини реалізується через сукупність білково-білкових взаємодій. Протягом останніх років запропоновано низку нових підходів і обчислювальних методів аналізу функціональних взаємодій між білками. Білково-білкові взаємодії мають велике значення у багатьох біологічних процесах — від формування клітинних макромолекулярних комплексів до складних процесів передачі сигналів, і тісно пов'язані з механізмами патогенезу багатьох хвороб, а також з підходами до їхнього лікування. Зважаючи на це, а також на доступність великої кількості повністю секвенованих геномів, багато наукових дослідницьких груп проводять інтенсивні дослідження білково-білкових взаємодій різних організмів [85—88], використовуючи переважно двогібридні системи [86, 87]. Проте експериментальне виявлення білково-білкових взаємодій за допомогою двогібридних систем, застосоване під час так званих wide-scale досліджень (тобто таких, що охоплюють всю або практично всю множину білків досліджуваного організму), має ряд недоліків [86, 88], а саме: велику кількість помилково-позитивних (false-positive) результатів, низьку кореляцію таких досліджень на одному організмі між даними різних дослідницьких груп, неможливість знаходження короткотривалих, нестійких взаємодій, необхідність окремо аналізувати кожну пару білково-білкових взаємодій. З урахуванням цього великого значення набуває попередній біоінформаційний аналіз даних, який дозволяє зосередити увагу на меншій кількості високоймовірних взаємодій та підвищити ефективність їхнього виявлення.

Вищенаведене пояснює інтенсифікацію розробки та застосування протягом останніх 3—4 років методів *in silico* пошуку та виявлення білково-білкових взаємодій, до яких належать, серед інших, метод філогенетичного профілювання [89, 90], генних кластерів [91, 92], «Розетського каменю» або «спаяних генів» [93—95], пошук за структурною гомологією [96] та передбачення на основі первинної структури білка [97].

З використанням декількох із цих підходів нами проаналізовано білково-білкові взаємодії тирозил-тРНК синтетази та знайдено ряд потенційних білків-партнерів, здатних формувати з цим ферментом локальну інтерактому (рис. 3) [98].

Дослідження білково-білкових взаємодій (білкові мережі або інтерактоми) проводиться для ряду модельних організмів, зокрема, нематоди *Caenorhabditis elegans* [99]. Для підмножини білків, специфічних для таксону *Metazoa*, вже ідентифіковано

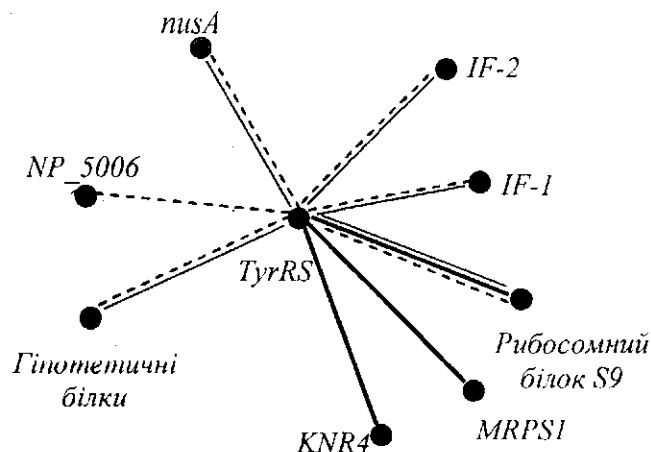


Рис. 3. Локальна інтерактома тирозил-тРНК синтетази. Білки-партнери було знайдено та передбачено за допомогою експериментальних (жирні лінії) та двох біоінформаційних підходів: методу генних кластерів (пунктирні лінії) і філогенетичного профілювання (тонкі лінії)

більш ніж 4000 взаємодій з використанням скринингування у дріжджовій двогібридній системі. Незалежний метод коафінного очищення в цілому узгоджується з даними двогібридної системи. Ще одним підходом є численні алгоритми комп'ютерного передбачення пар взаємодіючих білків, або «інтерологів». Сучасна версія карти інтерактоми нематоди (Worm Interactome Map, WI5) охоплює приблизно 5500 взаємодій. Аналіз топологічних і біологічних властивостей цих інтерактомичних мереж та їхня інтеграція з транскриптомними і протеомними даними надають можливість для численних біологічних передбачень, які, однак, вимагають експериментальної перевірки.

Широкомасштабні карти міжбілкових взаємодій відкривають нову перспективу для глобального аналізу функцій білків клітини та дозволяють простежити еволюцію живих систем на молекулярному рівні. Як приклад можна навести сукупні дані по інтерактомі дріжджів *S. cerevisiae* (рис. 4, див. вклейку). Іншим прикладом є база даних PSIMAP (Protein Structural Interactome Map) [100], у якій зібрано всі відомості про структурні взаємодії між надродинами білкових доменів з відомою просторовою структурою з банку даних PDB. У межах PSIMAP в єдину систему також систематизовано функціональну і еволюційну інформацію про ці білкові надродинами. Показано, що надродинами з високим ступенем відомих взаємодій, як правило, є дуже дивергентними у різних таксонів і, отже, належать до найдавніших за походженням. Виявле-

но, що надродина гідролаз нуклеотидтрифосфатів має найбільшу кількість взаємодій та є своєрідним центром усієї сітки міжбілкових взаємодій клітини.

**Молекулярна динаміка білків.** Функція білків на рівні інтрактому реалізується через їхню динаміку. Білки — це молекулярні машини, функція яких пов'язана з рухливістю у певних упорядкованих напрямках. Для розуміння механізму дії білків, як правило, недостатньо лише знання їхньої статичної кристалографічної структури. Саме тому вивчення конформаційної гнучкості білка є необхідною складовою процесу дослідження каталітичного механізму цього білка чи механізму його інгібування. Моделювання методом молекулярної динаміки (МД) є важливим інструментом вивчення конформаційної рухливості біологічних макромолекул, зокрема білків. За останні 10 років метод МД став популярним завдяки точності та якості отриманих результатів. Незважаючи на те, що симуляція методом МД є теоретичним підходом, його дані є вагомими, оскільки дослідники отримують інформацію про переміщення кожного атома системи в кожний момент часу.

Розрахунки методом МД можна проводити за допомогою пакетів програм Gromacs [102], NAMD, Moldy та ін. Динамічні моделі функціонування на сьогодні отримано для багатьох білків. Такі дані зберігаються на спеціалізованих серверах, наприклад, Morpheg, MolMovDB [103], ChemVis. За допомогою спеціальних програм (зокрема, VMD [104]) проводиться візуалізація розрахунків МД.

Аналіз методом МД зводиться до перебору дозволених конформацій у рамках певного енергетичного мінімуму, що дозволяє оцінити діапазон можливих змін у структурі білка. Розрахунки з використанням сучасних комп'ютерних програм з МД та персональних комп'ютерів усе ще обмежені часовим діапазоном у кілька наносекунд, однак застосування мультипроцесорних кластерів та нових версій програм дозволять перейти до розрахунку траєкторій в десятки і сотні наносекунд. Визначення значних конформаційних змін білків у мілісекундному діапазоні, які мають місце під час рухів окремих білкових доменів або у великих макромолекулярних ансамблях, потребує тривалих і складних комп'ютерних розрахунків та використання додаткових підходів.

Вивчення динаміки функціональних центрів білків методами моделювання МД суттєво розширює можливості структурної біології. В поєднанні з експериментальними методами це дозволяє перейти до аналізу не лише просторових, але й часових характеристик формування та функціонування складних надмолекулярних комплексів. Це сто-

сується як білково-білкових комплексів, так і комплексів білків з нуклеїновими кислотами, які утворюються на всіх етапах відтворення і реалізації генетичної інформації.

Дослідження конформаційної рухливості С-модуля тирозил-тРНК синтетази. Методом молекулярної динаміки нами проаналізовано С-модулі тирозил-тРНК синтетази бика і людини [52]. Виявлено добру кореляцію між конформаційною рухливістю в розчині для модельної структури та експериментально визначеної кристалографічної просторової структури цих модулів. Отримані результати задовільно узгоджуються із середньоквадратичними флуктуаціями  $C_{\alpha}$ -атомів С-модуля людини, розрахованими з урахуванням кристалографічних В-факторів. Показано, що найбільшу рухливість у розчині має поверхнева петля Val86-Lys92, яка бере участь у зв'язуванні тРНК [52].

**Молекулярна динаміка ВІЛ-1 протеази.** Одним з прикладів застосування методу МД є вивчення динаміки ВІЛ-1 протеази. ВІЛ-1 протеаза є ключовим білком процесингу поліпротеїнових попередників функціональних білків ВІЛ і тому була обрана як перспективна мішень для створення інгібіторів — ліків проти СНІДу. Для функціонування протеази необхідні значні конформаційні перебудови — відкриття «флепів», великих  $\beta$ -шпильок, які екранують активний центр ферменту і не дозволяють субстрату входити в активний центр. Отже, важлива роль конформаційної рухливості протеази для її функціонування є очевидною. Крім того, каталітична реакція і конформаційна стабільність ВІЛ-1 протеази залежать від рН середовища і концентрації іонів  $Na^{+}$ .

Рентгеноструктурні методи дослідження кристалічних структур ВІЛ протеази не можуть забезпечити детального описання механізмів конформаційних переходів та впливу середовища на динаміку білка. З використанням методу МД нами досліджено конформаційну рухливість ВІЛ-1 протеази в умовах, що найбільш відповідають фізіологічним [105—107].

З'ясовано, що конформаційна стабільність протеази при нейтральному значенні рН суттєво залежить від взаємодії каталітичних амінокислот Asp25/Asp25' з оточенням. При взаємодії каталітичної діади з катіонами середовища (іони  $Na^{+}$ ) відбувається екранування негативних зарядів каталітичних амінокислот і стабілізація просторової структури ВІЛ-1 протеази. Молекулярна динаміка протеази за таких умов має дуже подібний характер до МД ВІЛ-1 протеази, де каталітична діада є монопротонованою.

Слід зазначити, що в монопротонованому стані



активного центра ВІЛ-1 протеази є конформаційно найбільш стабільною та ензиматично активною. Існує альтернативний варіант взаємодії ВІЛ-1 протеази з оточенням при нейтральному рН — взаємодія з молекулою води. При цьому одна молекула води зв'язується з каталітичним центром таким чином, що це призводить до значної деформації четвертинної структури ВІЛ-1 протеази. Зниження стабільності цього ензиму при нейтральних рН ми пояснюємо конкуруючою реакцією зв'язування між іоном  $\text{Na}^+$  та молекулою води [106].

З використанням МД вивчали конформаційну гнучкість флепів, які утворюють верхню частину сайту зв'язування ліганду. Нами отримано такі конформації ВІЛ-1 протеази, де можуть утворюватися нові центри зв'язування для лігандів. Це відкриває шлях до створення інгібіторів, які могли б специфічно блокувати «динамічну» конформацію ВІЛ-1 протеази. Одержати подібні конформації для кристалічного стану ВІЛ-1 протеази рентгенографічними дослідженнями до цього часу не вдалося.

Одним із фундаментальних питань молекулярної біології є вплив мутацій на структуру і функції білків. Вплив дистальних мутацій амінокислотних залишків, які просторово віддалені від активного центра фермента, широко відомий, але експериментальними методами не завжди можна встановити цей взаємозв'язок. Висунувши робочу гіпотезу щодо впливу дистальних мутацій на динаміку білка і опосередковано через динаміку на активний центр ферменту, ми провели моделювання МД ВІЛ-1 протеази та її мутантів, резистентних до дії інгібіторів. На основі отриманих даних знайдено зміни в конформаційній рухливості протеази внаслідок цих мутацій. В результаті виявлено механізм впливу деяких мутацій на резистентність до інгібіторів.

Моделювання МД, хоч і є високоефективним засобом дослідження внутрішньомолекулярної рухливості білків, але, як і інші теоретичні методи, має свої обмеження, що можуть суттєво впливати на результати розрахунків. Для контролю «якості» розрахунків нами розроблено алгоритми порівняння даних МД з експериментальними даними ЯМР-спектроскопії. Зокрема, розроблено алгоритм, який дозволяє розраховувати параметр упорядкування структури  $S^2$  з даних МД та порівнювати його з даними ЯМР [106].

Моделювання взаємодій з лігандами. Якщо є доступними кристалографічні або модельні дані про будову активного центра ферменту або сайту зв'язування рецептора, то можна провести моделювання їхньої взаємодії з малими молекулами. При цьому аналізуються як стерична відповідність по-

верхонь макромолекули (рецептор) та малої молекули (ліганд), так і різні види міжмолекулярних взаємодій: електростатичні, ван-дер-ваальсові, гідрофобні, водневі зв'язки та ін. Як правило, використовують варіант докину з гнучким (flexible) лігандом і жорстким (rigid-body) рецептором, який більше відповідає можливостям сучасних комп'ютерів, але все значнішу увагу починають приділяти моделюванню конформаційної динаміки активного центра рецептора та взаємної зміни геометрії ліганду і рецептора. При цьому використовують різні варіанти комбінаторних та генетичних алгоритмів, а також методи Монте-Карло та молекулярної динаміки.

У більшості випадків проводять так званий первинний віртуальний скринінг значної кількості структур різних хімічних сполук з використанням бібліотек, які можуть включати сотні тисяч малих молекул, та обирають кілька десятків найкращих за енергетикою зв'язування потенційних лігандів. Далі аналізують можливі спільні риси таких сполук (пошук фармакофорів) та проводять комбінаторний або спрямований дизайн їхніх хімічних похідних, з якими проводять повторний раунд скринінгу.

Для точнішого розуміння взаємодії ліганд—рецептор здійснюють і розрахунки *ab initio* для зони контакту, однак останні вимагають значних обчислювальних потужностей. Описані розрахунки можна виконати, використовуючи безкоштовні пакети програм докину, які працюють під Linux: DOCK, AutoDock [107], Gold [108], Grid, ICM-Dock та ін. До комерційних пакетів належать FlexX [109], Amber [110], Sybyl, Insight II. Інший підхід до розробки потенційних інгібіторів методом проектування їхньої структури *de novo* втілено в таких програмах, як LUDI [111].

Хімічний синтез відібраних сполук і експериментальне визначення констант інгібування/зв'язування в умовах *in vitro*, а також визначення специфічності дії і загальної токсичності в експериментах *in vivo* є наступними кроками на цьому шляху. Кінцевою метою є створення нових лікарських препаратів (drug discovery або drug design), спрямованих на блокування певних функціональних груп в активному центрі макромолекули.

Інший напрямок полягає в моделюванні взаємодії двох макромолекул, як правило, білково-білкових і білково-нуклеїнових комплексів, що є дуже актуальним, оскільки в принципі може бути застосованим до великої кількості просторових структур макромолекул з метою пошуку ймовірних макромолекулярних комплексів і взаємодій в маштабах цілісної протеоми клітини. Основою біль-

шості методів є пошук геометричної комплементарності поверхонь двох макромолекул з включенням повного перебору їхніх взаємних орієнтацій. Наступним етапом є аналіз енергетики потенційних взаємодій та взаємної динамічної підгонки поверхонь контакту.

Серед програм, призначених для макромолекулярного докину, широко застосовують 3D-Dock (FTDock) [112], HEX [113], BIGGER [114], GRAMM [115], ZDOCK [116].

Комп'ютерний дизайн нових лікарських препаратів. Перехід до постгеномної ери в біології спричинив появу нових можливостей у створенні лікарських препаратів шляхом комп'ютерного дизайну специфічних інгібіторів ферментів. Важливо, що дослідник може визначити основні мішені для спрямованого дизайну інгібіторів білків, які є ключовими компонентами певного захворювання. Процес створення нового лікарського препарату можна розділити на окремі етапи: пошук гена-мішені в геномі; аналіз амінокислотної послідовності білка-мішені; моделювання просторої структури білка-мішені; візуалізація та аналіз будови активного центра; дизайн потенційного інгібітора; хімічний синтез та дослідження інгібітора *in vitro* та *in vivo*.

Якщо просторова структура білка відома, використовують прямі методи конструювання лікарського препарату: аналіз структури центра зв'язування інгібітора, молекулярний докинг ліганду в активному центрі, аналіз стеричної комплементарності та характер зв'язків у комплексі білок—ліганд і т. д. Але якщо просторову структуру білка-мішені не визначено, то методи комп'ютерного моделювання дозволяють побудувати таку модель, візуалізувати активний центр білка та провести наступні етапи дизайну нового інгібітора як потенційного лікарського препарату.

До недавнього часу кількість можливих мішеней для створення лікарських препаратів була обмеженою, тоді як тепер розвиток геноміки, протеоміки, транскриптоміки та інтерактоміки в комплексі з методами біоінформатики відкриває широкі перспективи для пошуку тисяч нових мішеней та розробки нових, більш ефективних і безпечних ліків.

Український біоінформатичний web-портал BioUA. Для проведення досліджень в галузі біоінформатики, геноміки та структурної біології в 2002—2003 рр. створено і введено в дію першу версію web-порталу як внутрішнього біоінформаційного ресурсу Інституту молекулярної біології і генетики НАН України (<http://www.bio.ua.org.ua>). На його основі планується розробити основний вітчизняний тримовний web-портал в галузі біоін-

форматики і структурної біології. Одним з основних ресурсів web-порталу BioUA є впорядкована система анованих лінків, які надають посилання на провідні світові біоінформаційні ресурси. Ці лінки систематизовано за відповідними тематичними розділами: 1) системи класифікації генів і генних родин; 2) ідентифікація і передбачення функцій генів та їхніх продуктів; 3) аналіз будови генів і регуляції транскрипції та альтернативного сплайсингу (транскриптоміка); 4) зв'язок генів зі спадковими та інфекційними хворобами; 5) аналіз експресії генів і відповідні технології; 6) аналіз промоторних ділянок генів; 7) методи аналізу, порівняння і моделювання просторової структури білків; 8) динаміка макромолекул; 9) віртуальний скринінг та дизайн *de novo* потенційних інгібіторів до активних центрів ферментів. Портал створюється як динамічна база даних, що дозволить здійснювати його неперервне оновлення і вдосконалення та сконцентрувати в одному місці значну кількість відібраної і анованої інформації в галузі дослідження молекулярних основ функціонування геному і його регуляції.

Планується створення на порталі спеціалізованих web-платформ, орієнтованих на практичне застосування та проведення спеціальних досліджень в галузі структурної біоінформатики: моделювання просторових структур нових білків-мішеней з геномів патогенних організмів, моделювання молекулярної динаміки білків тощо

Біоінформатика є однією з найдинамічніших галузей сучасної біологічної науки, що надзвичайно швидко розвивається. Її віднесено до одного з пріоритетних напрямків в Шостій рамочній програмі, яка координує європейські наукові дослідження, а також Національного наукового фонду США та інших країн. Слід зазначити, що розвиток біоінформатики став можливим завдяки успіхам геноміки та переходу біології у постгеномну еру. Біоінформатика є новим напрямком фундаментальної науки, яка дозволить найближчим часом суттєво розширити горизонти пізнання біологічних процесів на молекулярному рівні і створити біомедицину майбутнього.

K. A. Odynets, S. S. Ivahno, D. B. Kovalsky, B. T. Tokovenko, A. I. Kornelyuk

Structural bioinformatics in post-genomic era

Summary

Structural bioinformatics is a novel branch of biology which uses the computational methods of analysis with the aim of modeling the 3D structures of proteins and macromolecular complexes. The goal of structural bioinformatics research is also the creation of novel modulators of functional activities of proteins, such as novel drugs

(drug design). Progress in structural bioinformatics is determined by the rapid growth in the deciphering of several hundreds of genomes of both prokaryotic and eukaryotic organisms and transition to post-genomic era. In this review some aspects of structural bioinformatics such as 3D structure modeling, proteomics and interactomics, bioinformatics in transcriptome analysis, molecular dynamics of proteins, protein-ligand interactions modeling and computer aided drug design are discussed. The development and applications of these methods for mammalian tyrosyl-tRNA synthetase and viral HIV-protease are discussed. The first Ukrainian web-portal BioUA devoted to the genomics and structural bioinformatics investigations is described in this review.

К. А. Одынец, С. С. Ивахно, Д. Б. Ковальский, Б. Т. Токовенко, А. И. Корнелюк

Структурная биоинформатика в постгеномную эру

Резюме

Структурная биоинформатика — новая область биологии, использующая методы компьютерного анализа для моделирования пространственной структуры белков и макромолекулярных комплексов для выяснения их функций. Целью структурной биоинформатики является также создание селективных модуляторов функциональной активности белков как новых лекарственных препаратов. Прогресс структурной биоинформатики обусловлен бурным развитием геномики — расшифровкой нескольких сотен геномов про- и эукариотических организмов и переходом биологии в постгеномную эру. В обзоре рассмотрены такие вопросы, как моделирование пространственной структуры белков, протеомика и интерактомика, биоинформатика при анализе транскриптома, молекулярная динамика белков, моделирование взаимодействий с лигандами, компьютерный дизайн лекарственных препаратов. Описано применение этих подходов при исследовании тирозил-тРНК синтетазы млекопитающих и протеазы вируса иммунодефицита человека. Для проведения исследований в области геномики и структурной биоинформатики создан первый украинский биоинформационный web-портал BioUA.

#### ПЕРЕЛІК ЛІТЕРАТУРИ

- Orengo C. A., Jones D. T., Thornton J. M. Bioinformatics: genes, proteins and computers // BIOS.—Oxford, 2003.—298 p.
- Yu U., Lee S. H., Kim Y. J., Kim S. Bioinformatics in the post-genome era // J. Biochem. Mol. Biol.—2004.—37, N 1.—P. 75—82.
- Lindvall J. M., Blomberg K. E., Smith C. I. In silico tools for signal transduction research // Brief Bioinform.—2003.—4, N 4.—P. 315—324.
- Kellam P., Holzerlandt R., Gramoustianou E., Jenner R., Kwan A. Viral bioinformatics: computational views of host and pathogen // Novartis Found Symp.—2003.—254.—P. 234—252.
- Wu T. D. Bioinformatics in the post-genomic era // Trends Biotechnol.—2001.—19, N 12.—P. 479—480.
- Eisenberg D., Marcotte E. M., Xenarios I., Yeates T. O. Protein function in the post-genomic era // Nature.—2000.—405, N 6788.—P. 823—826.
- Pruitt K. D. WebWise: guide to the Sanger Center's web site // Genome Res.—1998.—8, N 1.—P. 4—8.
- Bernal A., Ear U., Kyrpides N. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide // Nucl. Acids Res.—2001.—29, N 1.—P. 126—127.
- Venter J. C., Adams M. D., Myers E. W., et al. The sequence of the human genome // Science.—2001.—291, N 5507.—P. 1304—1351.
- Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E. The Protein Data Bank // Nucl. Acids Res.—2000.—28, N 1.—P. 235—242.
- Wheeler D. L., Church D. M., Edgar R., Federhen S., Helmberg W., Madden T. L., Pontius J. U., Schuler G. D., Schriml L. M., Sequeira E., Suzek T. O., Tatusova T. A., Wagner L. Database resources of the National Center for Biotechnology Information: update // Nucl. Acids Res.—2004.—32.—P. D35—D40.
- Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D., Bairoch A. ExpASY: The proteomics server for in-depth protein knowledge and analysis // Nucl. Acids Res.—2003.—31, N 13.—P. 3784—3788.
- Benson D. A., Karsch-Mizrachi I., Lipman D. J., Ostell J., Wheeler D. L. GenBank: update // Nucl. Acids Res.—2004.—32.—P. D23—D26.
- Kulikova T., Aldebert P., Althorpe N., Baker W., Bates K., Browne P., van den Broek A., Cochrane G., Duggan K., Eberhardt R., Faruque N., Garcia-Pastor M., Harte N., Kanz C., Leinonen R., Lin Q., Lombard V., Lopez R., Mancuso R., McHale M., Nardone F., Silventoinen V., Stoehr P., Stoesser G., Tuli M. A., Tzouvara K., Vaughan R., Wu D., Zhu W., Arweiler R. The EMBL nucleotide sequence database // Nucl. Acids Res.—2004.—32.—P. D27—D30.
- Boeckmann B., Bairoch A., Arweiler R., Blatter M.-C., Estreicher A., Gasteiger E., Martin M. J., Michoud K., O'Donovan C., Phan I., Pilbout S., Schneider M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003 // Nucl. Acids Res.—2003.—31, N 1.—P. 365—370.
- Chen J., Anderson J. B., DeWeese-Scott C., Fedorova N. D., Geer L. Y., He S., Hurwitz D. J., Jackson J. D., Jacobs A. R., Lanczycki C. J., Liebert C. A., Liu C., Madej T., Marchler-Bauer A., Marchler G. H., Mazumder R., Nikolskaya A. N., Rao B. S., Panchenko A. R., Shoemaker B. A., Simonyan V., Song J. S., Thiessen P. A., Vasudevan S., Wang Y., Yamashita R. A., Yin J. J., Bryant S. H. MMDB: Entrez's 3D-structure database // Nucl. Acids Res.—2003.—31, N 1.—P. 474—477.
- Guex N., Peitsch M. C. Swiss-model and the Swiss-PdbViewer: An environment for comparative protein modeling // Electrophoresis.—1997.—18, N 15.—P. 2714—2723.
- Koradi R., Billeter M., Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures // J. Mol. Graph.—1996.—14, N 1.—P. 51—55.
- Martz E. Protein Explorer: easy yet powerful macromolecular visualization // Trends Biochem. Sci.—2002.—27, N 2.—P. 107—109.
- Sayle R. A., Milner-White E. J. RASMOL: biomolecular graphics for all // Trends Biochem. Sci.—1995.—20, N 9.—P. 374.
- Одынець К. О., Корнелюк О. І. Методи аналізу та моделювання просторової структури білків // Наукові записки Києво-Могилян. Академії.—2001.—19.—С. 7—17.
- Murzin A., Brenner S. E., Hubbard T., Chothia C. SCOP: a structural classification of proteins for the investigation of sequences and structures // J. Mol. Biol.—1995.—247, N 4.—P. 536—540.
- Orengo C. A., Michie A. D., Jones S., Jones D. T., Swindells M. B., Thornton J. M. CATH — a hierarchic classification of protein domain structures // Structure.—1997.—5, N 8.—P. 1093—1108.
- Shindyalov I. N., Bourne P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path // Protein Eng.—1998.—11, N 9.—P. 739—747.

25. Holm L., Sander C. Touring protein fold space with Dali/FSSP // Nucl. Acids Res.—1998.—26, N 1.—P. 316—319.
26. Dodge C., Schneider R., Sander C. The HSSP database of protein structure-sequence alignments and family profiles // Nucl. Acids Res.—1998.—26, N 1.—P. 313—315.
27. Schwede T., Kopp J., Guex N., Peitsch M. C. SWISS-MODEL: an automated protein homology-modeling server // Nucl. Acids Res.—2003.—31, N 13.—P. 3381—3385.
28. Bates P. A., Kelley L. A., MacCallum R. M., Sternberg M. J. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM // Proteins.—2001.—Suppl. 5.—P. 39—46.
29. Combet C., Jambon M., Deleage G., Geourjon C. Geno3D: automatic comparative molecular modelling of protein // Bioinformatics.—2002.—18, N 1.—P. 213—214.
30. Iwatake M., Kawakita S., Umeyama H. FAMS and FAMSBASE: modeling of all the genomes and database // Genome Inform.—2002.—13.—P. 565—566.
31. Marti-Renom M. A., Stuart A., Fiser A., Sanchez R., Melo F., Sali A. Comparative protein structure modeling of genes and genomes // Annu. Rev. Biophys. and Biomol. Struct.—2000.—29.—P. 291—325.
32. Vriend G. WHAT IF: a molecular modelling and drug design program // J. Mol. Graph.—1990.—8, N 1.—P. 52—56.
33. Canutescu A. A., Shelenkov A. A., Dunbrack R. L., Jr. A graph theory algorithm for protein side-chain prediction // Protein Sci.—2003.—12, N 9.—P. 2001—2014.
34. Kelley L. A., MacCallum R. M., Sternberg M. J. E. Enhanced genome annotation using structural profiles in the program 3D-PSSM // J. Mol. Biol.—2000.—299, N 2.—P. 499—520.
35. Rychlewski L., Jaroszewski L., Li W., Godzik A. Comparison of sequence profiles. Strategies for structural predictions using sequence information // Protein Sci.—2000.—9, N 2.—P. 232—241.
36. Shi J., Blundell T. L., Mizuguchi K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties // J. Mol. Biol.—2001.—310, N 1.—P. 243—257.
37. Meller J., Elber R. Linear programming optimization and a double statistical filter for protein threading protocols // Proteins.—2001.—45, N 3.—P. 241—261.
38. Bystroff C., Thorsson V., Baker D. HMMSTR: a hidden Markov model for local sequence-structure correlations in proteins // J. Mol. Biol.—2001.—301, N 1.—P. 173—190.
39. Melo F., Feytmans E. Assessing protein structures with a non-local atomic interaction energy // J. Mol. Biol.—1998.—277, N 5.—P. 1141—1152.
40. Colovos C., Yeates T. O. Verification of protein structures: patterns of nonbonded atomic interactions // Protein Sci.—1993.—2, N 9.—P. 1511—1519.
41. Laskowski R. A., MacArthur M. W., Moss D. S., Thornton J. M. PROCHECK: a program to check the stereochemical quality of protein structures // J. Appl. Cryst.—1993.—26.—P. 283—291.
42. Luthy R., Bowie J. U., Eisenberg D. Assessment of protein models with three-dimensional profiles // Nature.—1992.—356, N 6364.—P. 83—85.
43. Hoofst R. W. W., Vriend G., Sander C., Abola E. E. Errors in protein structures // Nature.—1996.—381, N 6580.—P. 272—282.
44. Johnson M. S., Lehtonen J. V. Comparison of protein three-dimensional structures // Bioinformatics: sequence, structure, and databanks: A practical approach / Eds D. Higgins, W. Taylor.—Oxford: Univ. press, 2002.—P. 15—50.
45. Taylor W. R., Orengo C. A. Protein structure alignment // J. Mol. Biol.—1989.—208, N 1.—P. 1—22.
46. Orengo C. A., Brown N. P., Orengo C. A. Fast structure alignment for protein database searching // Proteins.—1992.—14, N 2.—P. 139—167.
47. Orengo C. A., Jones D. T., Thornton J. M. Protein superfamilies and domain superfolds // Nature.—1994.—372, N 6507.—P. 631—634.
48. Madej T., Gibrat J.-F., Bryant S. H. Threading a database of protein cores // PROTEINS: Struct., Funct. and Genet.—1995.—23, N 3.—P. 356—369.
49. Russell R. B., Barton G. J. Multiple sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels // PROTEINS: Struct., Funct. and Genet.—1992.—14, N 2.—P. 309—323.
50. Голуб А. Г., Одынец К. А., Ньпорко А. Ю., Корнелюк А. И. Моделирование пространственной структуры СООН-концевого цитокин-подобного модуля цитоплазматической тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка.—2000.—16, № 6.—С. 515—524.
51. Одынец К. А., Вазылевский О. Е., Корнелюк О. І. Homology modeling of structure of NH<sub>2</sub>-terminal module of mammalian (*Bos taurus*) tyrosyl-tRNA synthetase // Биополимери і клітина.—2002.—18, № 6.—С. 547—550.
52. Каниболоцкий Д. С., Одынец К. А., Скурский С. И., Корнелюк А. И. Изучение внутримолекулярной подвижности цитокин-подобного С-концевого модуля тирозил-тРНК синтетазы млекопитающих методом молекулярной динамики // Физика живого.—2003.—11, № 2.—С. 61—71.
53. Корнелюк А. И., Курочкин И. В., Мацука Г. Х. Тирозил-тРНК синтетаза из печени быка. Выделение и физико-химические свойства // Молекуляр. биология.—1988.—22, № 1.—С. 176—186.
54. Гнатенко Д. В., Корнелюк А. И., Курочкин И. В., Мацука Г. Х. Высокомолекулярный лабильный комплекс тирозин-тРНК синтетазы из печени быка // Биополимеры и клетка.—1991.—7, № 1.—С. 63—69.
55. Гнатенко Д. В., Курочкин И. В., Рибкинська Т. А., Корнелюк А. И., Мацука Г. Х. Выделение и характеристика функционально активной протеолитически модифицированной формы тирозил-тРНК синтетазы из печени быка // Укр. биохим. журн.—1991.—63, № 4.—С. 61—67.
56. Клименко И. В., Корнелюк А. И., Мацука Г. Х. Конформационное изменение тирозил-тРНК-синтетазы из печени быка при взаимодействии с гомологичной тРНК по данным флуоресцентной спектроскопии // Биополимеры и клетка.—1993.—9, № 6.—С. 29—33.
57. Kornelyuk A. I., Klimenko I. V., Odynets K. A. Conformational change of mammalian tyrosyl-tRNA synthetase induced by tyrosyl adenylate formation // Biochem. and Mol. Biol. Int.—1995.—35, N 2.—P. 317—322.
58. Калачнюк Л. Г., Корнелюк О. І., Мацука Г. Х. Тирозинова тРНК (Q\*IA) печінки бика. Визначення ділянок її взаємодії з гомологічною аміноацил-тРНК синтетазою методом хімічної модифікації // Укр. биохим. журн.—1995.—67, № 5.—С. 60—65.
59. Леванец О. В., Найденов В. Г., Вудмаска М. И., Мацука Г. Х., Корнелюк А. И. ПЦР-амплификация, клонирование и секвенирование фрагмента кДНК, кодирующего нуклеотидсвязывающий домен тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка.—1996.—12, № 5.—С. 66—71.
60. Леванец О. В., Найденов В. Г., Вудмаска М. И., Мацука Г. Х., Корнелюк А. И. Клонирование кДНК, кодирующей С-концевую часть тирозил-тРНК синтетазы млекопитающих, с использованием ПЦР-амплифицированного ра-

- диоактивно меченного гибридизационного зонда // Биополимеры и клетка.—1997.—13, № 2.—С. 121—127.
61. Леванец О. В., Найденов В. Г., Одынец К. А., Вудмаска М. И., Мацука Г. Х., Корнелюк А. И. Гомология С-концевого некаталитического домена тирозил-тРНК синтетазы млекопитающих с цитокином ЕМАР II и некаталитическими доменами метионил- и фенилаланил-тРНК синтетаз // Биополимеры и клетка.—1997.—13, № 6.—С. 474—478.
  62. Дубровский А. Л., Савинская Л. А., Корнелюк А. И. Клонирование и бактериальная экспрессия цитокин-подобного некаталитического домена бычьей тирозил-тРНК синтетазы // Биополимеры и клетка.—1998.—14, № 5.—С. 449—452.
  63. Корнелюк А. И. Структурное и функциональное исследование тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка.—1998.—14, № 4.—С. 349—359.
  64. Корнелюк А. И. Белковая инженерия // Биополимеры и клетка.—2001.—17, № 6.—С. 520—532.
  65. Kornelyuk A. I., Tas M., Dubrovsky A., Murray J. C. Cytokine activity of the non-catalytic EMAP 2-like domain of mammalian tyrosyl-tRNA synthetase // Биополимеры и клетка.—1999.—15, № 2.—С. 168—172.
  66. Найденов В. Г., Вудмаска М. И., Корнелюк А. И., Мацука Г. Х. Сайт-направленный мутагенез остатков лизина, локализованных в соединительном пептиде нуклеотидсвязывающего домена (свертки Россмана) тирозил-тРНК синтетазы из печени быка // Биополимеры и клетка.—2000.—16, № 4.—С. 275—280.
  67. Froimowitz M. HyperChem: a software package for computational chemistry and molecular modeling // Biotechniques.—1993.—14, N 6.—P. 1010—1013.
  68. Hogue C. W. V. Structure databases // Bioinformatics: a practical guide to the analysis of genes and proteins / Eds A. D. Baxevanis, B. F. Francis-Ouellette.—London: John Wiley and Sons, 2001.—P. 83—107.
  69. Renault L., Kerjan P., Pasqualato S., Menetrey J., Robinson J. C., Kawaguchi S., Vassilyev D. G., Yokoyama S., Mirande M., Cherfils J. Structure of the EMAP II domain of human aminoacyl-tRNA synthetase complex reveals evolutionary dimer mimicry // EMBO J.—2001.—20, N 3.—P. 570—578.
  70. Yang L., Liu J., Skene R. J., Mcree D. E., Schimmel P. Crystal structure of an Emap-II-like cytokine released from a human tRNA synthetase // Helv. chim. acta.—2003.—86.—P. 1246.
  71. Kawaguchi S., Muller J., Linde D., Kuramitsu S., Shibata T., Inoue Y., Vassilyev D. G., Yokoyama S. The crystal structure of the tCsaA protein: an export-related chaperone from *Thermus thermophilus* // EMBO J.—2001.—20, N 3.—P. 562—569.
  72. Glaser F., Pupko T., Paz I., Bell R. E., Bechor-Shental D., Martz E., Ben-Tal N. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information // Bioinformatics.—2003.—19, N 1.—P. 163—164.
  73. Baldi P., Hatfield G. W. DNA microarrays and gene expression: From experiments to data analysis modeling.—Cambridge: Univ. press, 2002.
  74. Veculescu V. E., Zhang L., Vogelstein B., Kinzler K. W. Serial analysis of gene expression // Science.—1995.—270, N 5235.—P. 484—487.
  75. Schena M., Shalon D., Davis R. W., Brown P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray // Science.—1995.—270, N 5235.—P. 467—470.
  76. Heller M. J. DNA microarray technology: devices, systems, and applications // Annu. Rev. Biomed. Eng.—2002.—4.—P. 129—153.
  77. Fodor S. P., Read J. L., Pirran M. C., Stryer L., Lu A. T., Solas D. Light-directed, spatially addressable parallel chemical synthesis // Science.—1991.—251, N 4995.—P. 767—773.
  78. Lipshutz R. J., Fodor S. P., Gingeras T. R., Lockhart D. J. High density synthetic oligonucleotide arrays // Nat. Genet.—1999.—21, N 1.—P. 20—24.
  79. Basarsky T., Verdnic D., Zhai J. Y., Wellis D. Overview of a microarray scanner: design essentials for an integrated acquisition and analysis platform // Microarray biochip technology / Ed. M. Schena—New York: Eaton Publ., 2000.—P. 265—284.
  80. Alon U., Barkai N., Notterman D. A., Gish K., Ybarra S., Mack D., Levine A. J. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays // Proc. Nat. Acad. Sci. USA.—1999.—96, N 12.—P. 6745—6750.
  81. Miki R., Kadota K., Bono H., Mizuno Y., Tomaru Y., Carninci P., Itoh M., Shibata K., Kawai J., Konno H., Watanabe S., Sato K., Tokusumi Y., Kikuchi N., Ishii Y., Hamaguchi Y., Nishizuka I., Goto H., Nitanda H., Satomi S., Yoshiki A., Kusakabe M., DeRisi J. L., Eisen M. B., Iyer V. R., Brown P. O., Muramatsu M., Shimada H., Okazaki Y., Hayashizaki Y. Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18816 full-length enriched mouse cDNA arrays // Proc. Nat. Acad. Sci. USA.—2001.—98, N 5.—P. 2199—2204.
  82. Iyer V. R., Eisen M. B., Ross D. T., Schuler G., Moore T., Lee J. C., Trent J. M., Staudt L. M., Hudson J. Jr., Boguski M. S., Lashkari D., Shalon D., Botstein D., Brown P. O. The transcriptional program in the response of human fibroblasts to serum // Science.—1999.—283, N 5398.—P. 83—87.
  83. Quackenbush J. Computational genetics: computational analysis of microarray data // Nat. Rev. Genet.—2001.—2, N 6.—P. 418—427.
  84. Івахню С. С., Корнелюк О. І. Мікроарей: огляд технологій та аналіз даних // Укр. біохім. журн.—2004.—76, № 2.—С. 5—19.
  85. Dagkessamanskaia A., Martin-Yken H., Basmaji F. Interaction of Knr4 protein, a protein involved in cell wall synthesis, with tyrosine tRNA synthetase encoded by TYS1 in *Saccharomyces cerevisiae* // FEMS Microbiol. Lett.—2001.—200, N 1.—P. 53—58.
  86. Ito T., Chiba T., Ozawa R., Yoshida M., Hattori M., Sakaki Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome // Proc. Nat. Acad. Sci. USA.—2001.—98, N 8.—P. 4569—4574.
  87. Ito T., Tashiro K., Muta S., Ozawa R., Chiba T., Nishizawa M., Yamamoto K., Kuhara S., Sakaki Y. Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins // Proc. Nat. Acad. Sci. USA.—2000.—97, N 3.—P. 1143—1147.
  88. Uetz P., Giot L., Cagney G., Uetz P., Giot L., Cagney G., Mansfield T. A., Judson R. S., Knight J. R., Lockshon D., Narayan V., Srinivasan M., Pochart P., Qureshi-Emili A., Li Y., Godwin B., Conover D., Kalbfleisch T., Vijayadamar G., Yang M., Johnston M., Fields S., Rothberg J. M. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae* // Nature.—2000.—403, N 6770.—P. 623—627.
  89. Pellegrini M., Marcotte E. M., Thompson M. J., Eisenberg D., Yeates T. O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles // Proc. Nat. Acad. Sci. USA.—1999.—96, N 8.—P. 4285—4288.
  90. Tatusov R. L., Natale D. A., Garkavtsev I. V., Tatusova T. A., Shankavaram U. T., Rao B. S., Kiryutin B., Galperin M. Y., Fedorova N. D., Koonin E. V. The COG database: new

- developments in phylogenetic classification of proteins from complete genomes // *Nucl. Acid Res.*—2001.—29, N 1.—P. 22—28.
91. Bork P., Dandekar T., Diaz-Lazcoz Y., Eisenhaber F., Huynen M., Yuan Y. Predicting function: from genes to genomes and back // *J. Mol. Biol.*—1998.—283, N 4.—P. 707—725.
  92. Overbeek R., Fonstein M., D'Souza M., Pusch G. D., Maltsev N. The use of gene clusters to infer functional coupling // *Proc. Nat. Acad. Sci. USA.*—1999.—96, N 6.—P. 2896—2901.
  93. Enright A. J., Iliopoulos I., Kyrpides N. C., Ouzounis C. A. Protein interaction maps for complete genomes based on gene fusion events // *Nature.*—1999.—402, N 6757.—P. 86—90.
  94. Eisenberg D., Marcotte E. M., Xenarios I., Yeates T. Protein function in the post-genomic era // *Nature.*—2000.—405.—P. 823—826.
  95. Marcotte E. M., Pellegrini M., Ng H.-L., Rice D. W., Yeates T. O., Eisenberg D. Detecting protein function and protein-protein interactions from genome sequences // *Science.*—1999.—285, N 5428.—P. 751—753.
  96. Jaffe J., Robinson M., Bubeck D. Biophysics final project.—New York, 1999.
  97. Bock J. R., Gough D. A. Predicting protein-protein interactions from primary structure // *Bioinformatics.*—2001.—17, N 5.—P. 455—460.
  98. Токовенко Б. Т., Одинець К. О., Корнелюк О. І. Аналіз білок-білкових взаємодій тирозил-тРНК синтетаз методами біоінформатики // *Дні науки НАУКМА: Наук. записки НАУКМА.*—2003.—22, ч. III.—С. 376—379.
  99. Li S., Armstrong C. M., Bertin N., Ge H., Milstein S., Boxem M., Vidalain P. O., Han J. D., Chesneau A., Hao T., Goldberg D. S., Li N., Martinez M., Rual J. F., Lamesch P., Xu L., Tewari M., Wong S. L., Zhang L. V., Berriz G. F., Jacotot L., Vaglio P., Reboul J., Hirozane-Kishikawa T., Li Q., Gabel H. W., Elewa A., Baumgartner B., Rose D. J., Yu H., Bosak S., Sequerra R., Fraser A., Mango S. E., Saxton W. M., Strome S., Van Den Heuvel S., Piano F., Vandenhaute J., Sardet C., Gerstein M., Doucette-Stamm L., Gunsalus K. C., Harper J. W., Cusick M. E., Roth F. P., Hill D. E., Vidal M. A map of the interactome network of the metazoan *C. elegans* // *Science.*—2004.—303, N 5657.—P. 540—543.
  100. Bolser D., Dafas P., Harrington R., Park J., Schroeder M. Visualisation and graph-theoretic analysis of a large-scale protein structural interactome // *BMC Bioinformatics.*—2003.—4, N 1.—P. 45.
  101. Lindahl E., Hess B., van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis // *J. Mol. Model.*—2001.—7.—P. 306—317.
  102. Echols N., Milburn D., Gerstein M. MolMovDB: analysis and visualization of conformational change and structural flexibility // *Nucl. Acids Res.*—2003.—31, N 1.—P. 478—482.
  103. Humphrey W., Dalke A., Schulten K. VMD: visual molecular dynamics // *J. Mol. Graph.*—1996.—14, N 1.—P. 33—38.
  104. Ковальський Д. Б., Корнелюк А. І. Конформаційні зміни ВИЧ-1 протеази в області «флэпов»: вивчення методом молекулярної динаміки в піко- і наносекундному часових інтервалах // *Біополімери і клітина.*—2002.—18, № 2.—С. 117—123.
  105. Ковальський Д. Б., Каниболоцький Д. С., Дубина В. Н., Корнелюк А. І. Структурна подвижність ВИЧ-1 протеази: вплив протонірованості активного центру на конформацію ВИЧ-1 протеази в воді // *Укр. біохім. журн.*—2002.—74, № 6.—С. 124—128.
  106. Ковальський Д. Б., Іванова О. С., Дубина В. Н., Каниболоцький Д. С., Корнелюк А. І. Параметр упорядочення орієнтації N-H зв'язей як міра конформаційної подвижності білка: розробка алгоритма розрахунку з даних симуляції молекулярної динаміки і порівняння з даними ЯМР, показані на прикладі ВИЧ-1 протеази // *Укр. біохім. журн.*—2004.—75, № 3.
  107. Morris G. M., Goodsell D. S., Halliday R. S., Huey R., Hart W. E., Belew R. K., Olson A. J. Automated docking using Lamarckian genetic algorithm and an empirical binding free energy function // *J. Comp. Chem.*—1998.—19, N 14.—P. 1639—1662.
  108. Verdonk M. L., Cole J. C., Hartshorn M. J., Murray C. W., Taylor R. D. Improved protein-ligand docking using GOLD // *Proteins.*—2003.—52, N 4.—P. 609—623.
  109. Kramer B., Rarey M., Lengauer T. Evaluation of the FlexX incremental construction algorithm for protein-ligand docking // *PROTEINS: Struct., Funct. and Genet.*—1999.—37, N 2.—P. 228—241.
  110. Pearlman D. A., Case D. A., Caldwell J. W., Ross W. R., Cheatham T. E. III, DeBolt S., Ferguson D., Seibel G., Kollman P. AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules // *Comp. Phys. Commun.*—1995.—91.—P. 1—41.
  111. Bohm H. J. The computer program LUDI: a new method for the *de novo* design of enzyme inhibitors // *J. Comput. Aided. Mol. Des.*—1992.—6, N 1.—P. 61—78.
  112. Smith G. R., Sternberg M. J. Evaluation of the 3D-Dock protein docking suite in rounds 1 and 2 of the CAPRI blind trial // *Proteins.*—2003.—52, N 1.—P. 74—79.
  113. Ritchie D. W. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2 // *Proteins.*—2003.—52, N 1.—P. 98—106.
  114. Krippahl L., Moura J. J., Palma P. N. Modeling protein complexes with BiGGER // *Proteins.*—2003.—52, N 1.—P. 19—23.
  115. Vakser I. A. Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex // *Proteins.*—1997.—Suppl. 1.—P. 226—230.
  116. Chen R., Li L., Weng Z. ZDOCK: an initial-stage protein-docking algorithm // *Proteins.*—2003.—52, N 1.—P. 80—87.

УДК 577.322:577.332  
Надійшла до редакції 12.03.04