Вісн. Харк. ун-ту, № 606 БІОФІЗИЧНИЙ ВІСНИК Вип. 2 (13). 2003

МОЛЕКУЛЯРНА БІОФІЗИКА

УДК 541.144

СПРОЩЕНА МОДЕЛЬ ІЄРАРХІЧНОГО ФОЛДІНГУ БІЛКІВ З САМОУЗГОДЖЕНОЮ ЕНЕРГІЄЮ ЗВ'ЯЗКУ

С. О. Єсилевський, В. М. Харкянен

Інститут фізики НАН України, 03039 , Київ, Проспект науки 46.

Надійшла до редакції 18 листопада 2003 р.

В роботі розглядається спрощена модель фолдінгу білку, в якій білок розглядається як гомополімер, що компактизується ієрархічним чином. Вважається, що спочатку утворюються локальні кластери мономерів (кластери першого порядку), які потім об'єднуються у кластери вищих порядків. Утворення кластерів вищого порядку вважається набагато повільнішим за динаміку кластерів нижчого порядку і розглядається у адіабатичному наближенні. При утворенні кластерів в їх структурі "заморожуються" ступені вільності окремих мономерів, що веде до ефективного стискання доступного конформаційного простору ланцюгу і ефективного мінімуму на гіперповерхні вільної енергії, який є нативним станом ланцюгу. Енергія зв'язку мономерів у кластерах вважається залежною від крупномасштабної повільної узагальненої координати, яка характеризує ступінь сформованості нативної структури. В свою чергу ця координата залежить від середньої кількості утворених зв'язків, так, що система в цілому стає самоузгодженою. Показано, що така система може мати два дискретні стаціонарні стани – розгорнутий та нативний. Імовірності реалізації цих станів в залежності від температури та концентрації денатуранту у розчині повністю відповідають експериментальним даним.

КЛЮЧОВІ СЛОВА: фолдінг білків, ієрархічний фолдінг, одностадійна кінетика, самоузгоджена енергія взаємодії.

Проблема фолдінгу білків є одною з найактуальніших проблем сучасної біофізики [1,2]. Після того, як було секвеновано геноми багатьох організмів, в тому числі і людини, особливо гостро постала проблема визначення нативної тривимірної структури поліпептидних ланцюгів за їхньою відомою амінокислотною послідовністю. На сьогодні всі дослідження у галузі фолдінгу досить чітко розділилися на прикладні, покликані знайти практичне вирішення цієї проблеми для реальних білків, та ті, що не претендують на застосовність до реальних білків, проте покликані з'ясувати фундаментальні фізичні принципи фолдінгу [1-3]. Останні оперують різними типами спрощених моделей, які дозволяють відволіктися від надзвичайно складної організації конкретних поліпептидних ланцюгів і виявити їх універсальні фундаментальні властивості [1-4]. Спрощені моделі можна розділити на моделі з визначеною послідовністю та статистичні моделі. Перші можуть моделювати різні аспекти фолдінгу конкретних послідовностей і враховувати відмінності в амінокислотному складі білків, тоді як другі допускають аналітичні розв'язки і відображають загальну термодинаміку фолдінгу і статистичні властивості послідовностей, проте не здатні описати фолдінг конкретних модельних поліпептидів.

Статистичні моделі добре описують експериментально спостережуваний "перехід фолдінгу", який носить характер фазового переходу першого роду у скінченій системі [1-3]. При зміні керуючого параметра, яким може бути температура, концентрація денатуранту і т.п. відбувається кооперативний перехід ланцюгу від розгорнутого "клубка" до нативної конформації (або компактного стану для статистичних моделей). Проте всі існуючі на сьогодні статистичні моделі фолдінгу мають один суттєвий недолік – вони не дають вичерпного пояснення того експериментального факту, що реакція фолдінгу відбувається як одностадійний процес, тобто ланцюг перебуває або в нативному або в розгорнутому стані, тоді як концентрація можливих інтермедіатів нехтовно мала.

Останні експериментальні дослідження показали, що така поведінка притаманна не тільки ансамблю молекул у розчині а й окремій молекулі. Особливо показовим є експеримент описаний в роботі [5]. В ній було досліджено інгібітор хімотрипсину 2 (PWT CI2) за допомогою методу флуоресцентного резонансного переносу енергії (FRET) в одиночних молекулах. Молекули несли два флуоресцентних зонда розташовані в сайтах, що є просторово близькими в нативному стані глобули. При зміні концентрації денатуранту спостерігалося розгортання білку, яке супроводжувалося збільшенням відстані між мітками і чітко детектувалося по зміні сигналу FRET. Таким чином відстань між мітками була певною узагальненою координатою, що характеризував стан глобули. Було показано, що *одинична* молекула білку може існувати у

двох чітко розділених станах – розгорнутому та нативному. Імовірність реалізації цих станів залежить від концентрації денатуранту, але їх характеристики в термінах узагальненої координати є сталими. Така поведінка випливає з вигляду кривих вільної енергії вздовж узагальненої координати, які схематично наведені на Рис. 1.

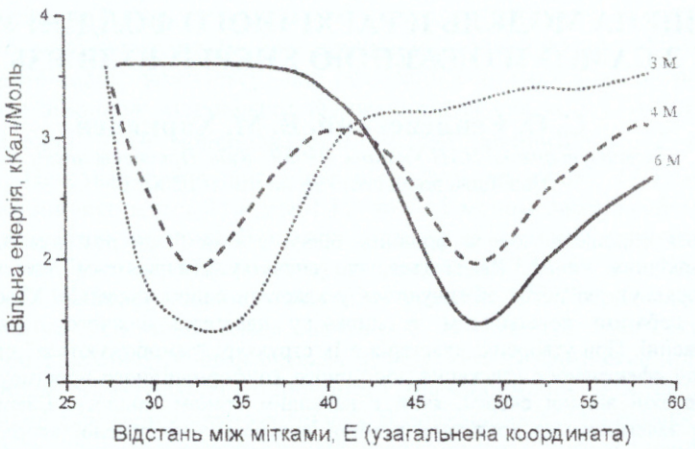

Рис. 1. Вільна енергія одиничної молекули білку PWT CI2 отримана методом FRET для різних концентрацій денатуранту (вказані поруч з кривими). Схематично відтворено з роботи [5].

В рамках розглянутого експерименту реакція фолдінгу окремої молекули є доланням досить високого потенціального бар'єру між денатурованим та нативним станами і відбувається як "класичний" Ейрінгівській активаційний одностадійний процес. Цим диктується одностадійність фолдінгу всього ансамблю молекул у розчині.

Експерименти такого роду однозначно показують, що будь яка теорія фолдінгу білків повинна враховувати не тільки енергетику та термодинаміку локальних парних контактів амінокислотних залишків, а й колективні внутрішні ступіні вільності глобули, що формується. Саме ці колективні ступені вільності відповідальні за істотно бістабільну поведінку окремих молекул і скоріш за все призводять до спостережуваної кінетики фолдінгу.

У попередніх роботах нами було запропоновано метод врахування колективної динаміки в процесі фолдінгу у граткових моделях, що досліджувалася кластер ним методом Монте-Карло [4,6]. В даній роботі ми розглядаємо спрощену статистичну модель фолдінгу яка враховує наявність в молекулі білку, що складається, колективних ступенів вільності. Модель дозволяє повністю відтворити експериментально спостережувані профілі вільної енергії і залежності імовірностей реалізації розгорнутого та нативного станів від концентрації денатуранту та температури та вичерпно пояснює одностадійний характер реакції фолдінгу. Характерною рисою моделі є явний опис звуження доступного конформаційного простору по мірі формування контактів між амінокислотними залишками ланцюгу, що складається.

МАТЕМАТИЧНА МОДЕЛЬ

Розглянемо модель гомополімеру, що складається з M мономерів з'єднаних абсолютно гнучкими і нерозтяжними однаковими зв'язками довжиною L. Об'ємами мономерів та зв'язків нехтуємо. Вважатимемо, що два мономера взаємодіють один з одним, коли вони знаходяться в межах певної малої відстані $l<<L$. За межами об'єму взаємодії рух мономерів вважатимемо вільною дифузією, що лімітується тільки довжиною зв'язків.

В такій моделі кожен мономер можна вважати певним обраним амінокислотним залишком, що може формувати енергетично вигідні контакти з іншими обраними залишками. Решта залишків при цьому формує гнучкі "петлі" і не бере активної участі в формуванні контактів.

Як вже було сказано пояснення новітніх експериментальних даних щодо фолдінгу одиночних молекул можливе лише при розгляді колективних рухів у глобулі, що формується. Відомо, що в білкових молекулах існує цілий спектр рухів з характерними часами від пікосекунд до хвилин [1-3]. З цього спектра необхідно виділити рухи, що є найбільш репрезентативними при розгляді різних стадій реакції фолдінгу, тобто мають

характерні часи співставні з середнім часом цих стадій. Таким чином необхідно ввести в модель певну ієрархію рухів з різними характерними часами. В даній роботі ієрархію рухів введено шляхом видозміни кластерної моделі фолдінгу, що була раніше розроблена авторами в контексті граткових моделей [4,6].

Згідно з цією моделлю в білку, що складається формуються кластери залишків, які з'єднані енергетично вигідними зв'язками. Кожен кластер поводиться як цілісний агрегат, що має власні поступальні та обертальні ступені вільності. Внутрішні ступені вільності окремих залишків в составі кластеру "заморожені", що призводить до сильного звуження конформаційного простору системи. Кластери взаємодіють один з одним формуючі кластери другого порядку, в яких в свою чергу "заморожені" ступені вільності кластерів першого порядку. Це знову звужує конформаційний простір системи. Процес "колапсу" конформаційного простору продовжується доки система не досягає унікального нативного стану. Детальні аргументи на користь цієї схеми та її докладний аналіз і застосування до конкретних граткових білків наведено в роботі [6]. Додамо до цієї схеми твердження про те, що формування кластерів вищого порядку відбувається набагато повільніше за формування кластерів попереднього порядку, тому швидші рухи можна розглядати в адіабатичному наближенні. "Заморожування" ступенів вільності нижчого порядку в такій схемі виникає цілком природно, оскільки вони формуватимуть певний профіль вільної енергії для рухів вищого порядку і не спостерігатимуться безпосередньо.

Спочатку розглянемо формування контакту між сусідніми мономерами ланцюгу i та i+1, яке може вважатися найшвидшим рухом в системі. Цю задачу можна розглядати як стандартну задачу сорбції: потрапляння одного мономеру ("ліганда") в зону взаємодії іншого. Ефективною концентрацією мономеру "ліганду" в першому наближенні є величина $C = \frac{\lambda}{L^3}$, де константа λ включає рухливість мономерів та величину зони взаємодії. Хай N_0 - імовірність того, що мономери знаходяться в контакті, тоді

$$\dot{N}_0 = C(1 - N_0) - hN_0,$$

де h – імовірність розриву зв'язку між мономерами.
В стаціонарному випадку маємо

$$N_0 = \frac{C}{C + h} = \frac{1}{1 + (h/\lambda) \cdot L^3}. \tag{1}$$

Тепер розглянемо формування кластера першого порядку в якому "злипаються" n сусідніх мономерів. Цей процес є набагато повільнішим за формування та руйнування окремого контакту і формує другий рівень часової ієрархії. Утворення кластера еквівалентне утворенню контакту між мономерами i та i+n. Очевидно, що ефективний об'єм, доступний для дифузії мономера "ліганда" i+n залежатиме від того чи сформовані індивідуальні найближчі контакти (i,i+1) та (i+1,i+2) і т.д. Оскільки формування кластеру відбувається набагато повільніше за динаміку локальних контактів, то ефективна концентрація "ліганду" становитиме

$$C_1 = \frac{\lambda}{\left(L[1 - N_0(i,i+1)] + L[1 - N_0(i+1,i+2)] + \ldots L[1 - N_0(n-1,n)]\right)^3} = \frac{\lambda}{\left(nL(1 - N_0)\right)^3}.$$

Коли жодний з контактів не сформований дифузія ліганду відбувається в об'ємі nL^3 , тоді як формування кожного з локальних контактів зменшує цей об'єм.

Аналогічно до (1) імовірність формування кластера першого рівня становитиме

$$N_1 = \frac{C_1}{C_1 + h} = \frac{1}{1 + (h/\lambda) \cdot [nL(1 - N_0)]^3}. \tag{2}$$

Розглянемо формування кластера другого порядку. Для спрощення вважатимемо, що в ньому об'єднуються також n кластерів першого порядку. Його формування еквівалентне утворенню контакту між залишками i та i+n^2 і відбувається набагато повільніше за динаміку кластерів першого порядку. Очевидно, що в формулі (2) достатньо замінити індекс 0 на 1, а формула (3) набуде вигляду

$$N_2 = \frac{C_2}{C_2 + h} = \frac{1}{1 + (h/\lambda) \cdot [n^2 L(1 - N_1)]^3}. \tag{3}$$

В загальному випадку кластера k-ого порядку маємо

$$N_k = \frac{C_k}{C_k + h} = \frac{1}{1 + (h/\lambda) \cdot [n^k L(1 - N_{k-1})]^3}, \tag{4}$$

що разом з (1) формує рекурентне співвідношення.

Визначимо імовірність відриву мономеру від кластеру h як активаційний процес подолання енергетичного бар'єру. Висота бар'єру дорівнює енергії взаємодії мономера з кластером, від якого він відривається H:

$$h = exp(-H/T),$$

де T – абсолютна температура в безрозмірних одиницях (вважаємо, що енергія вимірюється в одиницях k_BT).

В загальному випадку енергія відриву мономера від кластеру буде залежати від кількості мономерів у кластері, оскільки на кожен мономер буде певним чином діяти сумарне поле всіх інших. Тобто, енергія відриву буде залежати від ступеня сформованості нативної структури. Не уточнюючи деталей взаємодії введемо деяку додаткову енергію x, яка відображатиме цю залежність. Імовірність відриву перепишеться у вигляді

$$h = exp[-(H + x)/T],$$

а формула (4) набуде вигляду

$$N_k(H,x,T) = \frac{1}{1 + \left(exp[-(H+x)/T]/\lambda\right)\cdot[n^k L(1 - N_{k-1})]^3}. \qquad (5)$$

Постійна складова енергії сорбції H в свою чергу залежить від умов середовища. При введенні в систему денатуранту міцність зв'язків зменшується (зменшення H) і навпаки, при видаленні денатуранту міцність зв'язків зростає (збільшення H).

РЕЗУЛЬТАТИ Й ОБГОВОРЕННЯ

Перехід фолдінгу

Якщо покласти додаткову енергію $x = 0$, то модель перетворюється на модель колапсу гомополімеру з фіксованими енергіями парних зв'язків. Існує кількох подібних моделей, що використовують різні підходи до описання кооперативності процесу фолдінгу. В нашій моделі кооперативний характер утворення нативної структури закладений в самому визначенні ієрархічної кластеризації. Перехід фолдінгу, що відбувається при зміні температури або концентрації детануранту в нашій моделі показано на Рис. 2 та 3 для різних значень кількості ієрархічних рівнів M. Значення $M = 0$ відповідає окремому ізольованому контакту.

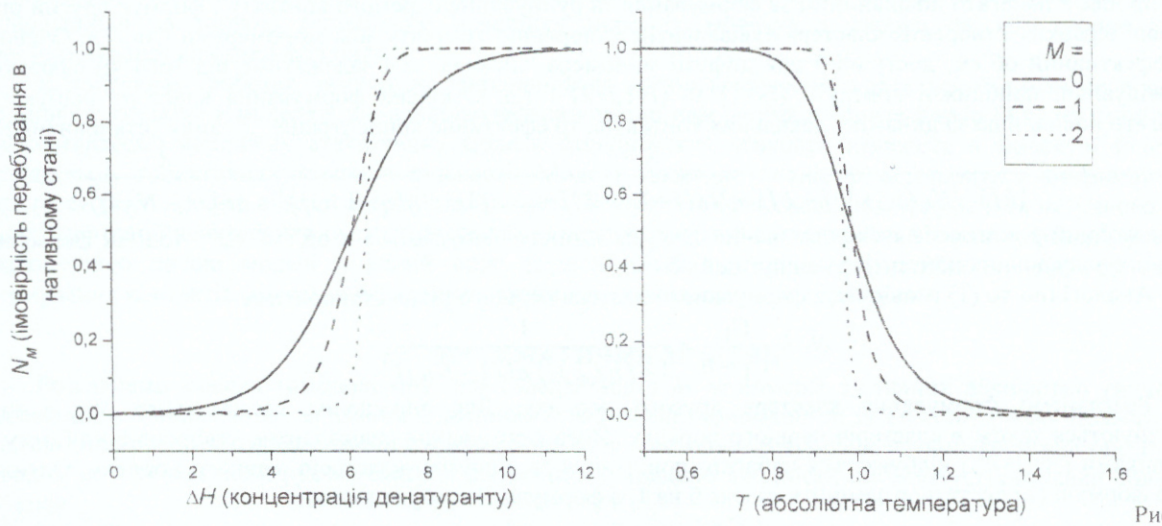

Рис. 2.

Перехід фолдінгу пр змінах енергії взаємодії мономерів ΔH (а) та підвищенні температури (б) для різних кількостей ієрархічних рівнів M у нативному білку. Параметри моделі: $H = 10$, $n = 2$, $L = 1$, $\lambda = 10^{-7}$, $x = 0$.

Видно, що при збільшенні кількості ієрархічних рівнів кооперативність переходу різко зростає, а його напівширина зменшується. У повній відповідності до експериментальних даних при зменшення H, що відповідає підвищенню концентрації денатуранту, та при підвищенні температури білок розгортається.

Структурна координата і бістабільність на рівні окремої молекули

Згідно з описаною моделлю динаміка кластерів попереднього ієрархічного рівня є набагато швидшою за динаміку кластерів наступного рівня, тому в системі існує чітка ієрархія часів характерних структурних перебудов. На практиці ступінь сформованості структури білку під час фолдінгу визначатиметься тими перебудовами, що мають характерні часи співставні з середнім часом фолдінгу, а отже є повільними по відношенню до більшості внутрішніх ступенів вільності. В залежності від експериментальної методики можуть фіксуватися різні зміни структури тому в загальному випадку можна розглянути деяку узагальнену конформаційну координату, рух вздовж якої розглядається як індикатор процесу фолдінгу. Розглянемо як впливає на цю координату динаміка швидких змінних системи.

Індивідуальні контакти у вже сформованій частині структури постійно формуються і руйнуються під дією теплових флуктуацій, але на великих часових масштабах для кожного з них існує певна імовірність перебування в "зв'язаному" стані. За час існування кожного індивідуального контакту вже сформована структура зазнає невеликої зміни, яка не встигне повністю релаксувати до нового стохастичного утворення цього контакту. Таким чином повільна структурна координата "відчуватиме" саме середню імовірність існування контакту. Як вже зазначалося, в загальному випадку енергія відриву індивідуального мономеру чи кластеру в свою чергу залежить від поточного стану структури, тобто від структурної координати. Таким чином утворюється самоузгоджена система, подібна до систем в яких існує зарядово-конформаційна взаємодія чи ліганд-структурна взаємодія [7-9]. Як було показано у численних попередніх роботах [7-9] таким системам проявляє властивості самоорганізації і має складну бістабільну поведінку.

Перейдемо до визначення структурної координати. Найбільш природнім на нашу думку є обрати структурною координатою додаткову енергію зв'язку E, яка повністю відповідає її якісному визначенню, наведеному вище. Такий вибір робить структурну координату незалежною від деталей взаємодій в системі та її конкретної просторової конфігурації, а отже є найбільш загальним.

Як повільна ступінь вільності, яка відповідає великомасштабним перебудовам системи, структурна координата скоріш за все матиме дифузійну стохастичну динаміку, яку можна описати рівнянням Ланжевена

$$\dot{x} = -\frac{\partial U(x)}{\partial x} + F + \sqrt{2D}\xi(t), \tag{6}$$

де $U(x)$ – структурний потенціал, який задає енергію системи зі значенням структурної координати x коли система "заморожена" на певному етапі процесу фолдінгу, тобто коли утворення і розпад контактів не відбувається; F – узагальнена сила, яка діє на структурну координату з боку стохастично виникаючих та зникаючих контактів; D – коефіцієнт дифузії вздовж структурної координати; $\xi(t)$ – нормально розподілена випадкова величина.

В найпростішому випадку можна вважати, що всі стохастичні зв'язки впливають на структурну координату адитивно, а вплив кожного є лінійним. В такому випадку можна записати

$$F(x) = \alpha \sum_{i=0}^{M} n^{M-i} \cdot N_i(H, x, T),$$

де α – емпірична константа. Ми врахували, що кількість контактів порядку i становить n^{M-i}.

В тому ж найпростішому випадку можна вважати структурний потенціал гармонічним:

$$\frac{\partial U(x)}{\partial x} = \frac{\beta}{2} \cdot x^2,$$

де β – константа. Слід зазначити, що реальні потенціали для повільних ступеней вільності в білку є суттєво негармонічними, але для нашої абстрактної моделі таке спрощення не призводить до якісного спотворення результатів. Отже остаточне рівняння для структурної координати набуває вигляду

$$\dot{x} = -\beta \cdot x + \alpha \sum_{i=0}^{M} n^{M-i} \cdot N_i(H, x, T) + \sqrt{2D}\xi(t) \tag{7}$$

У стаціонарному випадку (без урахування стохастичної складової) рівняння (7) запишеться як

$$\gamma \cdot x = \sum_{i=0}^{M} n^{M-i} N_i(H, x, T), \tag{8}$$

де $\gamma = \beta / \alpha$.

Це рівняння задає стаціонарні стани системи. Графічну інтерпретацію рівняння (8) наведено на Рис. 3.

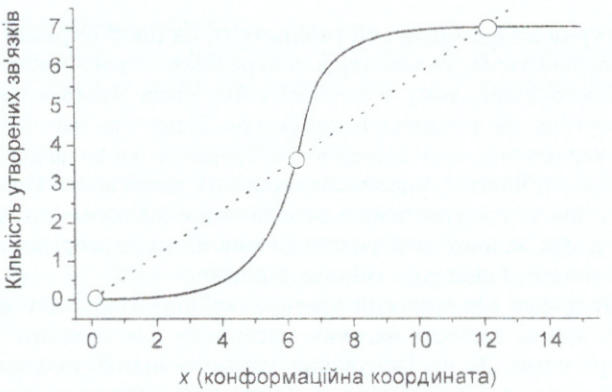

Рис. 3. Знаходження стаціонарних станів самоузгодженої системи (позначені колами) при різних температурі переходу фолдінгу $T = 1$. Крива відповідає правій частині рівняння (8), пряма – лівій.
Параметри моделі: $H = 10$, $n = 2$, $L = 1$, $\lambda = 10^{-7}$, $M = 2$.

Видно, що при певному підборі параметрів існує три стаціонарні стани, середній з яких є нестабільним, а два інших – стабільними [9]. Таким чином в системі можуть виникати бістабільні режими, в яких відбуваються стохастичні переключення між двома дискретними станами. Перший з них (лівий) відповідає малим значенням структурної координати і відповідно дуже малій кількості контактів в системі. Це розгорнутий стан ланцюгу. Другий (правий) відповідає великим значенням структурної координати та практично максимально можливій кількості контактів в системі. Це нативний стан повністю згорнутого білка. Переходи між цими станами можна описати ввівши *конформаційний потенціал* у вигляді

$$\frac{\partial U_{conf}(x)}{\partial x} = \int_{x_0}^{x} \left(\gamma \cdot x' - \sum_{i=0}^{M} n^{M-i} N_i(H, x', T) \right) dx',$$

де x_0 – початок відліку.

Конформаційний потенціал задає енергію самоузгодженої системи, що перебуває у рівновазі, при певному значенні структурної координати. Слід наголосити на відмінності конформаційного потенціалу від структурного. Структурний потенціал не враховує впливу стохастично виникаючих контактів, а описує лише "автономну", структурно обумовлену динаміку координати x. Конформаційний потенціал є вільною енергією самоузгодженої системи контактів та структурної координати.

Вигляд конформаційних потенціалів при різних значеннях зовнішніх параметрів наведено на Рис. 4.

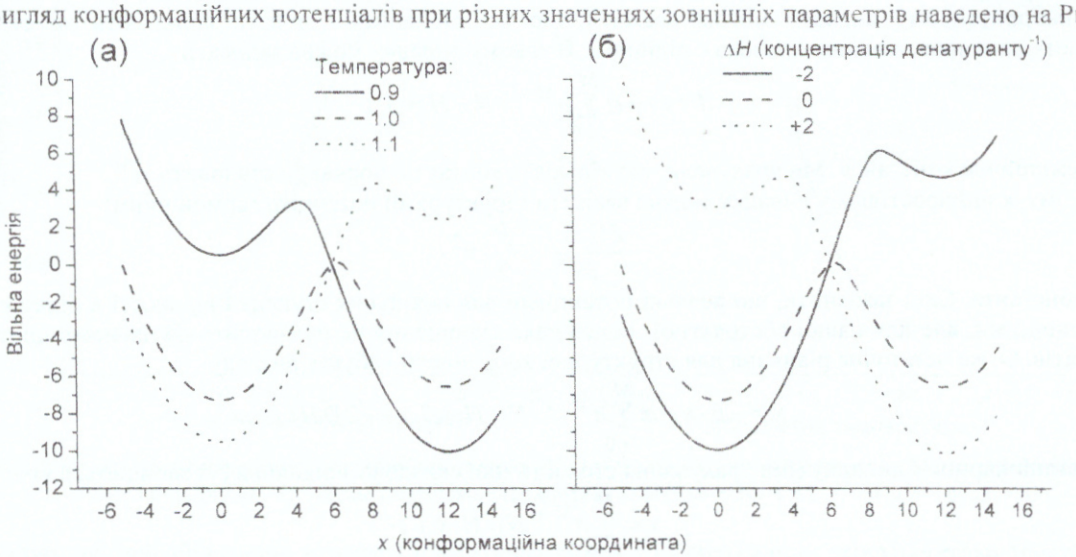

Рис. 4. Конформаційні потенціали самоузгодженої системи при різних температурах, що знаходяться поблизу температури переходу фолдінгу $T = 1$ (а) та при різних значеннях енергії зв'язку ΔH при $T = 1$.
Параметри моделі: $H = 10$, $n = 2$, $L = 1$, $\lambda = 10^{-7}$, $M = 2$.

Ліва потенціальна яма конформаційного потенціалу відповідає розгорнутій конформації, а права – нативній. Видно, що при зміні температури поблизу температури переходу фолдінгу відбувається різка зміна глибин ям. При підвищенні температури починає домінувати яма розгорнутого стану, при зменшенні – нативного. Аналогічна картина спостерігається при зміні параметра H, що моделює зміну концентрації денатуранта в розчині. При зменшенні H (збільшення концентрації денатуранту) відбувається перехід в розгорнутий стан, при збільшенні H (зменшення концентрації денатуранту) – перехід в нативний стан.

Порівняння Рис. 1 та 4б показує їх повну якісну ідентичність. Це свідчить про те, що запропонована модель повністю відтворює зафіксовану в експерименті форму профілю вільної енергії індивідуальної молекули білку, що згортається і дозволяє прослідкувати зміни цього профілю при зміні концентрації денатуранту в розчині.

ВИСНОВКИ

В роботі запропоновано модель фолдінгу білку із самоузгодженою енергією взаємодії. Модель базується на двох припущеннях: 1) наявності ієрархії часів характерних рухів у молекулі, що складається; 2) узгодженості енергії зв'язків між залишками та конформаційним станом ланцюгу.

Модель самоузгодженої енергії зв'язку дозволяє зрозуміти природу виникнення двох дискретних станів молекули – розгорнутого та нативного з боку загальних універсальних фізичних принципів, які не залежать від конкретної послідовності та топологічної організації білка в процесі фолдінгу. Наявність бістабільної поведінки на рівні однієї молекули дозволяє вичерпно пояснити спостережуваний одностадійний характер кінетики фолдінгу переважної більшості одно доменних білків. Модель дозволяє отримати профілі вільної енергії ланцюгу вздовж деякої узагальненої структурної координати, яку можна ототожнити зі структурно обумовленою координатою, що фіксується в сучасних експериментах на одиночних молекулах. Зміна форми профілю при зміні температури та концентрації денатуранту в розчині повністю відповідає експериментально зафіксованим залежностям. Це дозволяє зробити висновок, що концепція ієрархії часів характерних рухів та самоузгодженості енергій зв'язків, що покладені в основу моделі, імовірно мають місце в реальних білках.

СПИСОК ЛІТЕРАТУРИ

1. Sali A., Shakhnovich E., Karplus M. How does a protein fold? // Nature. – 1994. - № 369 (6477). – P.248- 251.
2. Dobson C.M., Sali A., Karplus M. Protein Folding: A Perspective from Theory and Experiment // Angew. Chem. Int. Ed. – 1998. V. 37. – P.868-893.
3. Mirny L., Shakhnovich E.. Protein folding theory: from lattice to all-atom models // Annu. Rev. Biophys. Biomol. Struct. - 2001. – V. 30. – P.361-396.
4. Yesylevskyy S.O. Demchenko A.P. Modeling the hierarchical protein folding using clustering Monte-Carlo algorithm. // Protein and Peptide Letters. – 2001. – V. 6. – P.437-442 .
5. Michalet X. and Weiss S. Single-molecule spectroscopy and microscopy // C. R. Physique. – 2002. – V. 3. – P.619-644.
6. Yesylevskyy S.O. and Demchenko A.P. Towards realistic description of collective motions in the lattice protein folding models // J. Biophys. Chem. - 2003. - In press.
7. Goushcha A.O., Kharkyanen V.N., Scott G.W., Holzwarth A.R. Self-Regulation Phenomena in Bacterial Reaction Centers. I. General Theory // Biophys. J. – 2000. – V. 79. – P.1237-1252.
8. Yesylevskyy S.O. and Kharkyanen V.N. Hierarchy of motions and quasi-particles in the simplified model of potassium channel selectivity filter // J. Biological Physics. - 2003. - In press.
9. Демченко А.П., Устинов А.А., Харкянен В.Н. Аналитическая теория лиганд-индуцированных конформационных переходов в белках // Вісник Харківського національного університету ім. В.Н. Каразіна. - № 525. - Біофізичний вісник. – 2001. - Вип. 1(8). - С.7-17.