

Аналіз вторинної структури цитокін-подібного С-кінцевого модуля тирозил-тРНК синтетази ссавців (*Bos taurus*) методами кругового дихроїзму та комп'ютерного моделювання

К.О. Одинець*, О.І. Корнелюк

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

Резюме. СООН-кінцевий модуль цитоплазматичної тирозил-тРНК синтетази (TyrRS) ссавців є допоміжним тРНК-зв'язувальним доменом, який також виявляє цитокін-подібну активність в ізольованому стані після протеолітичного відщеплення від каталітичного кору ферменту. У роботі проведено порівняльний аналіз вторинної структури цього біфункціонального С-модуля з TyrRS *Bos taurus* за допомогою експериментальних і теоретичних методів. Експериментальне визначення вторинної структури здійснено методом аналізу спектрів кругового дихроїзму (КД) препаратів рекомбінантного С-модуля. Теоретичне передбачення проведено 1) з використанням програм визначення вторинної структури С-модуля, виходячи з його амінокислотної послідовності, та 2) алгоритмів виведення вторинної структури з атомних координат (DSSP, STRIDE та P-SEA) просторової моделі С-модуля, побудованої за гомологією до цитокіну ЕМАР II людини. Експериментальні дані КД для С-модуля підтверджують нативний стан білка YCD3k в отриманих препаратах.

Ключові слова: тирозил-тРНК синтетаза, С-кінцевий модуль, передбачення вторинної структури білків, спектри кругового дихроїзму.

Вступ. Тирозил-тРНК синтетази (TyrRS) каталізують високоспецифічне приєднання L-тироzinу до гомологічних тРНК^{Tyr} і цим реалізують відповідні правила генетичного коду при функціонуванні білоксинтезуючої підсистеми клітин [1]. Усі відомі TyrRS є гомодимерами і можуть бути віднесені до двох типів — прокаріотичного (TyrRS еубактерій, мітохондрій і хлоропластів) та еукаріотичного (цитоплазма клітин еукаріотів і архебактерії). Послідовність цитоплазматичної TyrRS бика (*Bos taurus*) та її СООН-кінцевого модуля, який відповідає ділянці Val363-Ser528, була раніше визначена нами з її кДНК [2, 3].

Кожна з двох ідентичних субодиниць складається з 528 амінокислотних залишків (а.з.), які формують NH₂-кінцевий каталітичний та додатковий СООН-кінцевий модулі, з'єднані міжмодульним лінкером (рис. 1, А і Б). С-модуль має спорідненість до РНК, зокрема, до тРНК, а також виконує додаткову цитокін ЕМАР II-подібну функцію після його вивільнення при обмеженому протеолітичному розщепленні TyrRS *in vitro* та *in vivo* [4, 5]. Схожі ЕМАР II-подібні модулі присутні в усіх відомих амінокислотних послідовностях TyrRS хребетних та комах (таксон *Metazoa*), у гомологів білка р43 вищих еукаріотів, білка Arc1p дріжджів, а також у метіоніл-тРНК синтетазах рослин і хробаків (рівень гомології з С-модулем TyrRS становить 46—53 % ідентичності). У свою чергу, С-модуль складається з еволюційно давнього

* Corresponding author.

Tel.: +38044-5265589

E-mail address: odynets@imbg.org.ua

© К.О. Одинець, О.І. Корнелюк, 2005

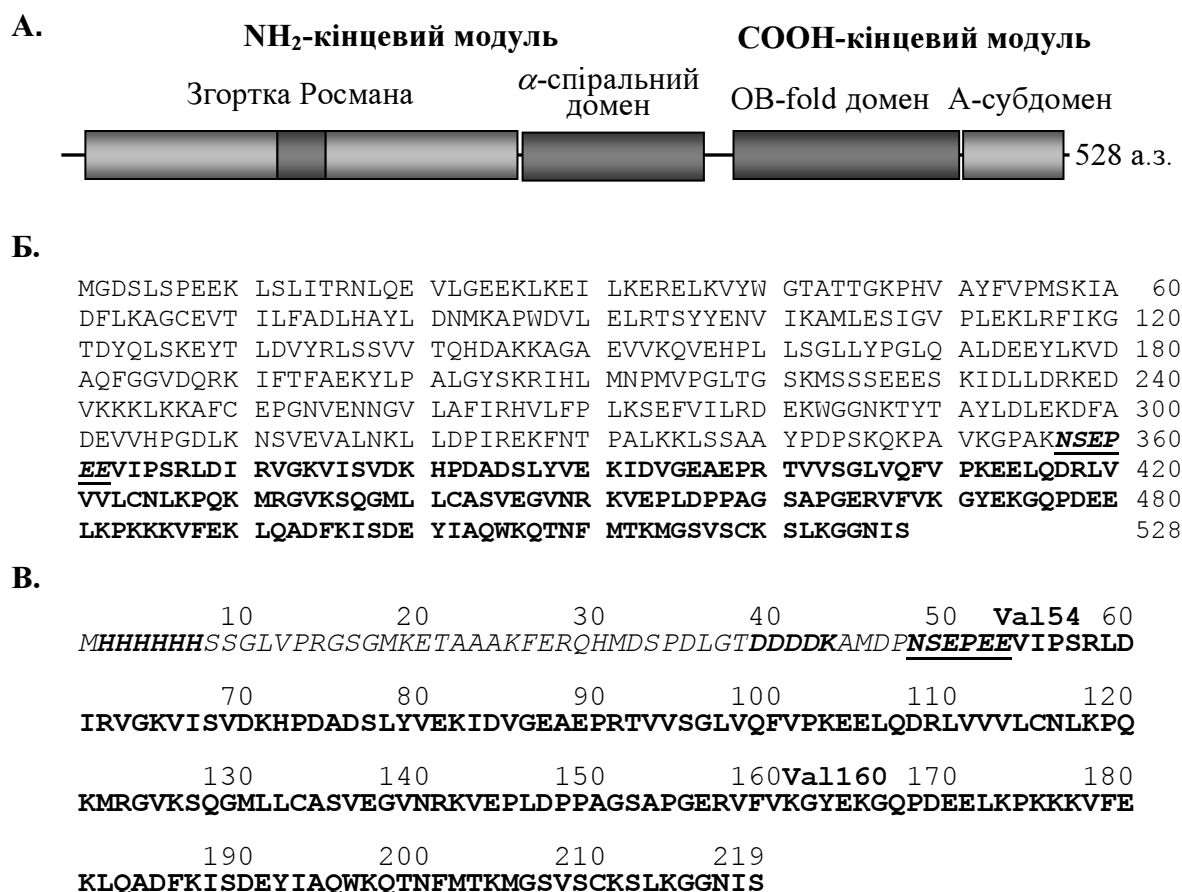


Рис. 1. **А.** Схема доменної організації однієї субодиниці цитоплазматичної тирозил-тРНК синтетази бика (*Bos taurus*). **Б.** Амінокислотна послідовність TyrRS бика, у складі якої C-модуль (Val363-Ser528) виділено напівжирним шрифтом. **В.** Амінокислотна послідовність рекомбінантного білка YCD3k, в якій виділено ділянку C-модуля, що включає домен, який має просторову згортку типу OB-fold (Val54-Phe159), та додатковий A-субдомен (Val160-Ser219).

домену з типом згортки OB-fold, який має переважно β -складчасту структуру (Val363-Phe468 у C-модулі TyrRS ссавців), та приєднаного до нього A-субдомену з вторинною структурою типу $\alpha+\beta$ (Val469-Ser528) [3]. Білкові домени зі згорткою типу OB-fold широко розповсюджені в усіх організмів і входять до складу білків, які взаємодіють з нуклеїновими кислотами. A-субдомен знайдено лише у складі ЕМАР II-подібних модулів, його просторова структура має унікальний тип згортання.

Кодувальну послідовність ЕМАР II-подібного C-модуля TyrRS бика було клоновано нами у векторі рЕТ-30а, що експресує рекомбінантний білок YCD3k завдовжки 219 а.з. [4, 5]. Цей білок містить із NH₂-кінця додаткові 45 а.з., що кодуються вектором, та невеликий фрагмент міжмодульного лінкера TyrRS (ділянка Asn357-Glu362, рис. 1В). N-кінцева послідовність містить (His)₆-послідовність, призначену

для афінної очистки білка, та сайт розщеплення ентерокиназою DDDDK[^]. У системі бактеріальної експресії нами було отримано гомогенні препарати C-модуля у значних кількостях, що дало змогу визначити його вторинну структуру методом аналізу спектрів кругового дихроїзму (КД-спектрів).

Нами також проводилося теоретичне визначення вторинної структури C-модуля і рекомбінантного білка YCD3k з їхніх амінокислотних послідовностей із застосуванням 11-ти спеціалізованих програм, які використовують різні алгоритми передбачення. З допомогою методу моделювання за гомологією було побудовано дві моделі просторової структури C-модуля, що ґрунтуються на експериментально визначених кристалографічних структурах гомологічних білків — домені В2 еубактеріальної фенілаланіл-тРНК синтетази (23 % ідентичності) [6] та цитокіні ЕМАР II людини

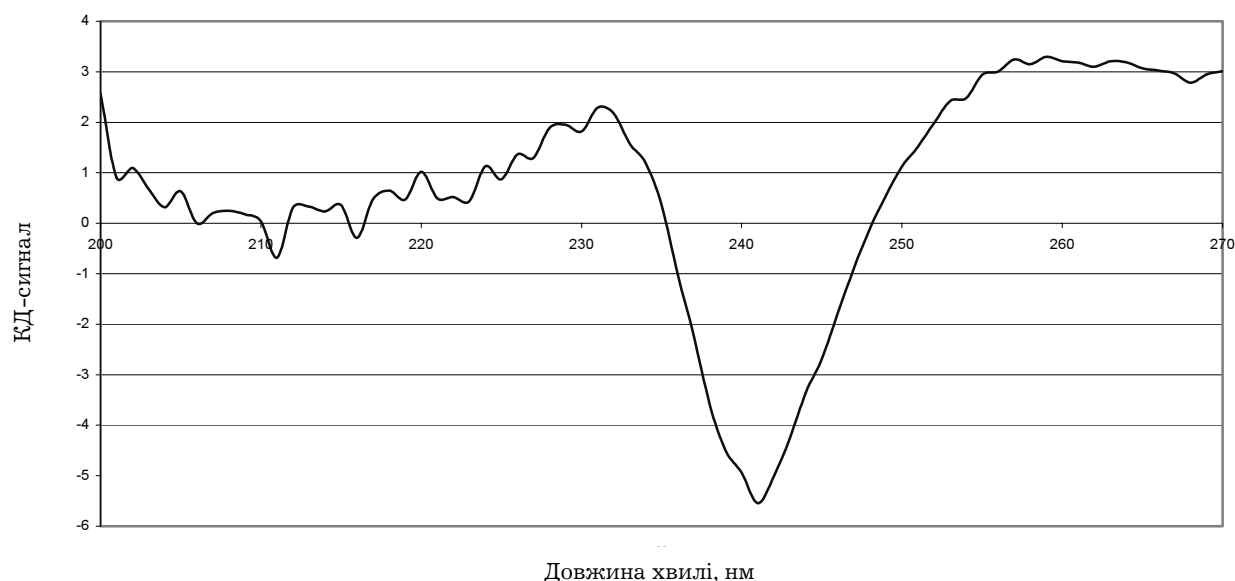


Рис. 2. Спектр кругового дихроїзму (КД) препарату рекомбінантного білка YCD3k, який містить С-модуль тирозил-тРНК синтетази бика і додаткові NH_2 -кінцеві 45 а.з., що кодує вектор рЕТ-30а.

(52,7 % ідентичності) [7]. З атомних координат моделі тривимірної структури С-модуля, побудованої за гомологією до цитокину ЕМАР ІІ людини, виведено вторинну структуру з використанням трьох різних алгоритмів. Отримані експериментальні й теоретичні дані порівнювалися між собою, а також зіставлялися із вторинною структурою С-модуля TyrRS людини, виведеною з рентгеноструктурних даних [8].

Матеріали і методи. Амінокислотну послідовність цитоплазматичної TyrRS бика було визначено раніше і депоновано до банків даних GenBank/GenPept (код AAC82467) і SwissProt (код Q29465). Основні фізико-хімічні властивості С-модуля і рекомбінантного білка YCD3k передбачено з використанням програми ProtParam із серверу ExPASy (<http://us.exPASy.org/tools/protparam.html>). Пошук рекомбінантних білків із визначеною структурою, які містять кодовану вектором рЕТ-30а послідовність, проведено з допомогою серверів NCBI і Protein Data Bank (PDB).

Отримання препаратів білка YCD3k. Рекомбінантний С-модуль TyrRS бика (білок YCD3k) одержували шляхом експресії у бактеріальній системі відповідної плазмідної конструкції рЕТ-30а/YCD3k у 400 мл суспензії клітин *Escherichia coli* штаму BL21(DE3) у середовищі LB при температурі 37 °С. Після отримання суспензії з оптичною густиною $A_{600} = 0,8$ проводили індукцію транскрипції гена

цільового білка, додаючи 0,5 мМ IPTG. Після 4—5 годин інкубації клітини осаджували, осад заморожували і зберігали при температурі -20 °С. Цей осад ресуспендували при температурі 0 °С у 15 мл буферного розчину (30 мМ трис- HCl , рН 8,0; 400 мМ NaCl, 5 мМ імідазолу) і проводили лізис клітин на ультразвуковому сонікаторі п'ять разів по 10 с. Отриманий клітинний лізат центрифугували при 6000 г упродовж 20 хв., супернатант наносили на Ni-NTA-агарозну колонку об'ємом 1,0 мл, урівноважену тим самим буферним розчином. Далі проводили ретельну промивку колонки 25 мл того ж буфера з додаванням 600 мМ NaCl та наступну ступінчасту елюцію буферами з додаванням 50, 100 і 250 мМ імідазолу. Фракції, які, за даними SDS-гель-електрофорезу, містили чистий цільовий білок (чистота > 95 %) і мали флуоресценцію більше 5 флуоресцентних одиниць при довжинах хвиль збудження 280 нм і випромінювання 335 нм, об'єднували і проводили їх діаліз проти 30 мМ Na-фосфатного буфера, рН 8,0; 200 мМ NaCl упродовж 12 годин.

Вимірювання й аналіз КД-спектрів. Спектри кругового дихроїзму препаратів рекомбінантного С-модуля вимірювали на дихрографі Aviv Circular Dichroism Spectrometer, Model 202 (Aviv, LakeWood N.J., США) під час термостатування при температурі 25 °С. Для проведення аналізу КД-спектрів використовували

Таблиця 1

Результати розрахунку вмісту основних елементів вторинної структури в рекомбінантному білку YCD3k, отримані з КД-спектра з використанням веб-серверу програми K2d (ALPHA — > 0,04 (4,0 %), BETA — > 0,48 (48,0 %), RANDOM — > 0,48 (48,0 %), SQUARE DISTANCE: 1046,07)

Довжина хвилі, нм	$\Delta \varepsilon$		Довжина хвилі, нм	$\Delta \varepsilon$	
	Експериментальні	Скореговані		Експериментальні	Скореговані
200	2,56	8,28	221	0,50	-6,02
201	0,90	7,11	222	0,52	-6,20
202	1,10	6,05	223	0,44	-6,41
203	0,68	4,94	224	1,13	-6,50
204	0,32	3,94	225	0,87	-6,47
205	0,63	3,01	226	1,36	-6,27
206	0,00	2,19	227	1,29	-5,98
207	0,20	1,50	228	1,88	-5,61
208	0,25	0,49	229	1,95	-5,10
209	0,18	-0,27	230	1,82	-4,48
210	0,04	-0,90	231	2,29	-3,81
211	-0,68	-1,62	232	2,18	-3,12
212	0,31	-2,33	233	1,58	-2,62
213	0,33	-2,93	234	1,19	-2,17
214	0,23	-3,49	235	0,36	-1,84
215	0,36	-3,94	236	-1,01	-1,53
216	-0,29	-4,34	237	-2,22	-1,28
217	0,48	-4,76	238	-3,58	-1,06
218	0,65	-5,14	239	-4,50	-0,90
219	0,47	-5,47	240	-4,93	-0,77
220	1,02	-5,77			

веб-сервер програми K2d (<http://www.embl-heidelberg.de/~andrade/k2d.html>) [9, 10], який дає змогу визначати вміст трьох основних типів вторинної структури — α -спіралей, β -тяжів і нерегулярної конформації. Для відповідних значень величин КД-спектрів рекомбінантного С-модуля в УФ-діапазоні від 200 до 240 нм (заданих як град \times см² \times дмоль⁻¹ помножене на 10⁻³) отримували передбачення процентного вмісту цих типів вторинної структури, а також оцінку точності їх визначення.

Передбачення вторинної структури з амінокислотної послідовності. Вторинну структуру рекомбінантного білка YCD3k, а також його фрагмента Val54-Ser219, що відповідає С-модулю TyrRS, було визначено із застосуванням мультипрограми метасерверу NPS@ (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_auto-mat.pl?page=/NPSA/npsa_seccons.html) [11], який поєднує 11 незалежних програм передбачення вторинної структури за різними алгоритмами, а саме: DPM, DSC, GOR1, GOR2, GOR4, HNNC, PHD, Predator, SIMPA96, SOPM, SOPMA. Незалежно використовували

веб-сервер програми PHD, який дає змогу одержувати більш детальну інформацію [12].

Моделювання просторової структури та виведення вторинної структури С-модуля.

Модель просторової структури С-модуля TyrRS білка було отримано методом моделювання за гомологією на основі експериментально визначеної просторової структури гомологічного білка ЕМАР ІІ людини (PDB коди 1EUJ, 1FL0, 1E7Z) [13, 14]. Вирівнювання амінокислотних послідовностей цільового білка і білка-шаблону проводили з використанням програми Swiss-PDB Viewer (DeepView) версії 3.7 (<http://us.expasy.org/spdbv/>) [15] з урахуванням результатів множинного вирівнювання всіх гомологічних послідовностей для більш точної локалізації делецій і вставок. Моделювання структури С-модуля здійснювалося з допомогою веб-серверу Swiss-Model версії 3.5 (<http://www.expasy.org/swiss-mod/SWISS-MODEL.html>) [15] та наступної процедури оптимізації петель і повної структури за методом молекулярної динаміки, описаним у роботі [7].

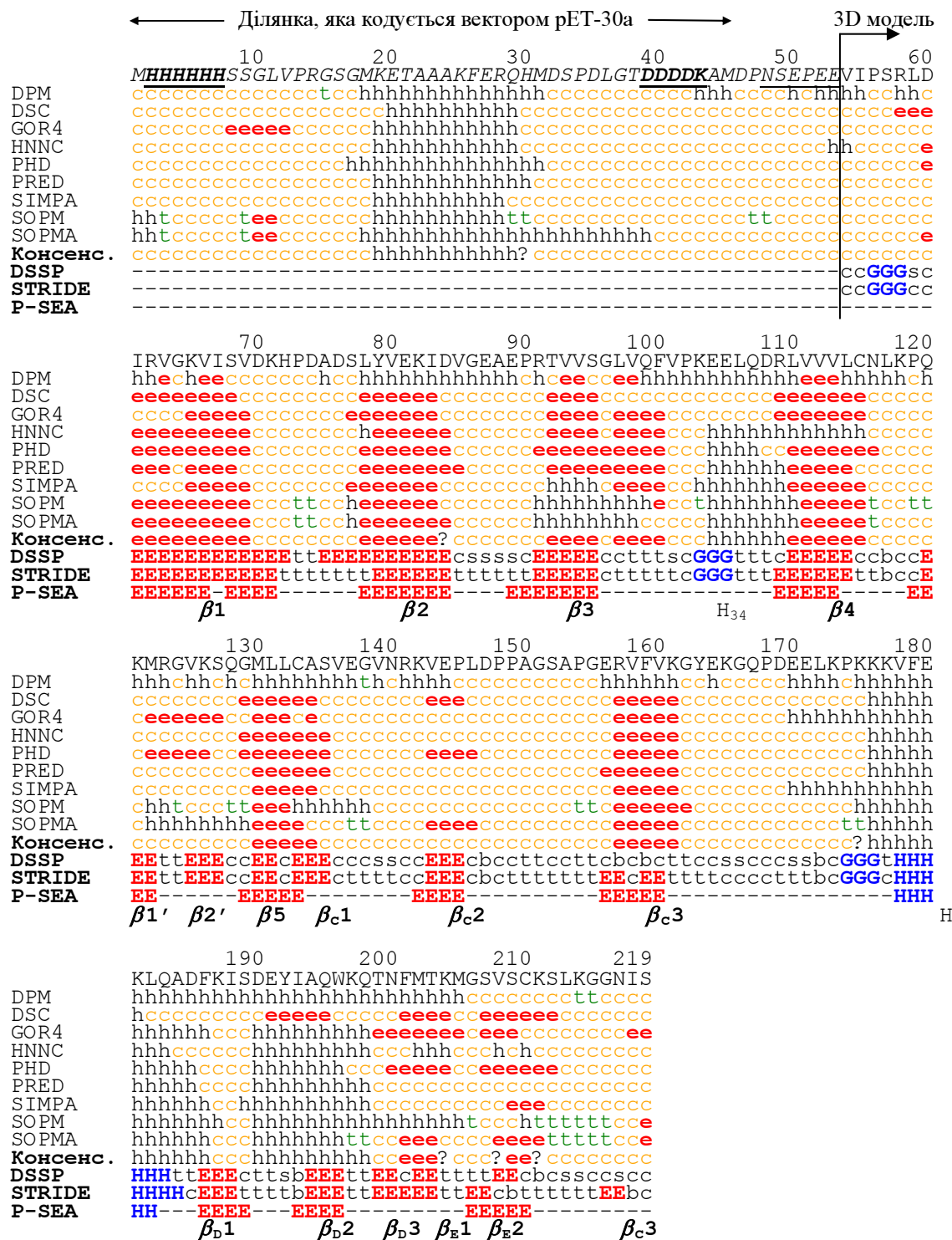


Рис. 3. Передбачення елементів вторинної структури з амінокислотної послідовності рекомбінантного білка YCD3k та виведення їх із просторових координат атомів моделі С-модуля. Наведено результати дев'яти з одинадцяти використаних програм, а також їхнє консенсусне передбачення. Стани (конформації) залишків подано в однобуквенному кодї (e – β-тяжі; h – α-спіралі; g – Z₁₀-спіралі; i – π-спіралі; b – β-містки; t – β-повороти; s – вигнуті ділянки; c – нерегулярна конформація; ? – незрозумілі стани; «-» – інші стани). Модель просторової структури С-модуля (ділянка Val54-Ser219) відповідає останнім 166 а.з. білка YCD3k і відповідно TyrRS. Виведення елементів вторинної структури з атомних координат моделі проведено за алгоритмами DSSP, STRIDE і P-SEA.

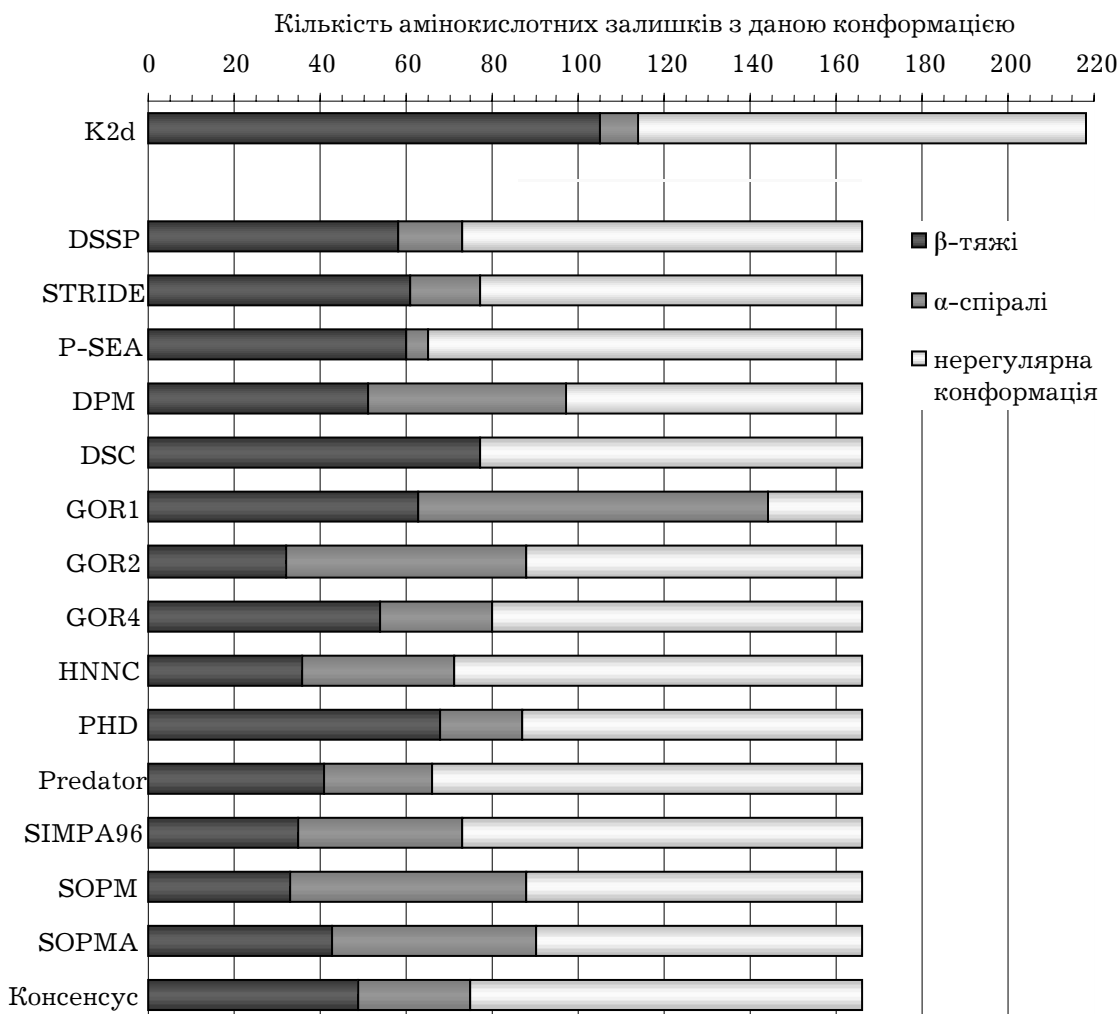


Рис. 4. Порівняння експериментальних даних, передбачення і виведення вторинної структури для ділянки рекомбінантного білка YCD3k Val54-Ser219, яка відповідає С-модулю TyrRS бика. Вказано кількість амінокислотних залишків, які знаходяться в одному з трьох станів (конформацій): β-тяжі, α-спіралі та нерегулярна конформація. Наведено результати: 1) аналізу КД-спектрів за допомогою програми K2d; 2) виведення елементів вторинної структури з атомних координат моделі просторової структури С-модуля з використанням алгоритмів DSSP, STRIDE і P-SEA; 3) передбачення вторинної структури з допомогою 11-ти програм та їх консенсусне передбачення.

Локалізацію і точні межі елементів вторинної структури С-модуля (ділянка Val363-Ser528 у TyrRS) виводили з атомних координат отриманої моделі просторової структури із застосуванням трьох різних алгоритмів: DSSP (<http://bioweb.pasteur.fr/seqanal/interfaces/dssp.html>) [16], STRIDE (<http://bioweb.pasteur.fr/seqanal/interfaces/stride.html>) [17] і P-SEA (http://bio-serv.cbs.cnrs.fr/HTML_BIO/frame_sea.html) [18]. Ці програми також використовували для аналізу вторинної структури гомологічного С-модуля TyrRS людини (PDB код 1NTG) [8].

Результати й обговорення. Основні фізико-хімічні властивості білка YCD3k. За даними програми ProtParam, рекомбінантний білок

YCD3k, який містить 219 а.з., має молекулярну масу 24255,6 кДа й теоретичне значення ізоелектричної точки $pI = 6,25$. Він складається з 3403 атомів, його молярний коефіцієнт екстинкції при 280 нм становить $9650 \text{ см}^{-1} \cdot \text{моль}^{-1}$, а оптичне поглинання розчину з концентрацією 1 мг/мл — 0,398 оптичних одиниць. Програма ProtParam класифікує цей білок як «стабільний» із передбаченим часом напівжиття >10 год. у клітинах *E. coli*.

Отримання й аналіз КД-спектрів. Для експериментального визначення кількісного вмісту основних типів вторинної структури білків широко використовується метод аналізу їхніх спектрів кругового дихроїзму [19–23].

Передбачення вмісту трьох основних конформацій вторинної структури у складі рекомбінантного білка YCD3k та власне С-модуля TugRS на основі їх амінокислотних послідовностей

Програма передбачення	Передбачений вміст елементів вторинної структури, а.з. (та %)*					
	Рекомбінантний білок YCD3k, 219 а.з.			С-модуль TugRS окремо, 166 а.з.		
	β -тяжі	α -спіралі	Нерегулярна конформація	β -тяжі	α -спіралі	Нерегулярна конформація
DPM	33 (15,1 %)	62 (28,4 %)	123 (56,4 %)	51 (30,7 %)	46 (27,7 %)	69 (41,6 %)
DSC	63 (28,9 %)	15 (6,9 %)	140 (64,2 %)	77 (46,4 %)	0 (0 %)	89 (53,6 %)
GOR1	60 (27,5 %)	118 (54,1 %)	40 (18,4 %)	63 (37,9 %)	81 (48,8 %)	22 (13,3 %)
GOR2	39 (17,9 %)	76 (34,9 %)	103 (47,3 %)	32 (19,3 %)	56 (33,7 %)	78 (47,0 %)
GOR4	57 (26,2 %)	37 (16,9 %)	124 (56,9 %)	54 (32,5 %)	26 (15,7 %)	86 (51,8 %)
HNNC	36 (16,5 %)	48 (22,0 %)	134 (61,5 %)	36 (21,7 %)	35 (21,1 %)	95 (57,2 %)
PHD	55 (25,2 %)	42 (19,3 %)	121 (55,5 %)	68 (40,9 %)	19 (11,5 %)	79 (47,6 %)
Predator	41 (18,8 %)	37 (16,9 %)	140 (64,2 %)	41 (24,7 %)	25 (15,1 %)	100 (60,2 %)
SIMPA96	35 (16,1 %)	48 (22,0 %)	135 (61,9 %)	35 (21,1 %)	38 (22,9 %)	93 (56,0 %)
SOPM	33 (15,1 %)	65 (29,8 %)	120 (55,1 %)	33 (19,9 %)	55 (33,1 %)	78 (46,9 %)
SOPMA	42 (19,3 %)	74 (33,9 %)	102 (46,8 %)	43 (25,9 %)	47 (28,3 %)	76 (45,8 %)
Консенсус	45 (20,6 %)	39 (17,9 %)	134 (61,5 %)	49 (29,5 %)	26 (15,7 %)	91 (54,8 %)
Діапазон, %	15,1–28,9 %	6,9–54,1 %	18,4–64,2 %	19,3–46,4 %	0–48,8 %	13,3–60,2 %

*Примітка. Вказано кількість амінокислотних залишків та їх відсоток від загальної кількості залишків (219 і 166 а.з. відповідно), β -повороти включено до нерегулярної конформації.

Створено низку програм і веб-серверів для аналізу КД-спектрів, зокрема, застосований нами веб-сервер програми K2d [9, 10], який використовує алгоритм нейронної сітки, самооптимізованої для виведення вмісту елементів вторинної структури з множини КД-спектрів для білків із референтної бази даних, вимірних в УФ-діапазоні 200–240 нм. Ця інформація зберігається як матриця результатів від сотень незалежних запусків цієї нейронної сітки. Також надається ймовірна помилка передбачення, яка базується на статистиці процедури навчання нейронної сітки.

Вимірювання КД-спектрів у діапазоні 200–270 нм проводили в 30 мМ натрій-фосфатного буфера, рН 8,0; 200 мМ NaCl при температурі 25 °С. Типовий отриманий спектр подано на рис. 2. Розрахований за даними програми K2d вміст β -тяжів у структурі рекомбінантного С-модуля становить 48,0 % (тобто 105 із 219 а.з.), α -спіралей — 4,0 % (9 а.з.), нерегулярних конформацій — 48,0 % (104 а.з.) (див. табл. 1).

Передбачення вторинної структури з амінокислотної послідовності. Вторинна структура рекомбінантного білка YCD3k та окремо його ділянки Val54-Ser219 визначалася з допомогою мультипрограмного веб-серве-

ру Network Protein Sequence @analysis (NPS@) [11], який поєднує 11 незалежних програм, а також надає їхнє усереднене консенсусне передбачення. У табл. 2, на рис. 3 і 4 для білка YCD3k подано сумовані результати цих програм для вмісту елементів вторинної структури за трьома основними станами (типами конформації) — β -тяжі, α -спіралі й нерегулярна конформація (враховуючи β -повороти). Можна констатувати досить значні розходження між даними різних програм, які вкладаються в діапазони 15,1–28,9 % для β -тяжів і 6,9–54,1 % для α -спіралей.

Окремо проводився аналіз для фрагмента Val54-Ser219 завдовжки 166 а.з., для якого з метою порівняння результатів різних методів було побудовано модель просторової структури. На рис. 3 показано, що більшість елементів вторинної структури розпізнається всіма програмами, але є значні розбіжності між даними передбачення для β -тяжів (19,3–46,6 %), α -спіралей (0–48,8 %) і нерегулярної конформації (13,3–60,2 %) (табл. 2). Слід зазначити, що програма PHD, яка враховує еволюційні дані з множинного вирівнювання гомологічних послідовностей [12], за даними порівняльного тестування точніше розпізнає межі елементів вторинної структури і коректно визначає

близько 80 % усіх амінокислотних залишків в одній із трьох основних конформацій. Тому дані саме цієї програми (40,9 % β -тяжів, 11,5 % α -спіралей і 47,6 % «інші стани») є *a priori* більш надійними, ніж консенсус за всіма 11-ма програмами, тому їх можна взяти за основу. З іншого боку, результати застарілих алгоритмів GOR1 і GOR2 значно відрізняються від усереднених.

Моделювання за гомологією просторової структури С-модуля та виведення з неї вторинної структури. Просторову структуру перших 53 а.з. рекомбінантного білка YCD3k не було змодельовано через відсутність для них білка-шаблону з визначеною просторовою структурою і достатнім рівнем гомології. Для штучно введеної послідовності Met1-Met45, що кодується вектором рЕТ-30а, проведено пошук за базою даних Protein Data Bank серед інших рекомбінантних білків із визначеною просторовою структурою, які містять цей самий структурний елемент. Серед чотирьох знайдених білків, які містять повні ідентичні послідовності (*Acetohydroxyacid synthase* (PDB код 1N0H), *Botulinum neurotoxin* (1E1H), *ССА-Adding enzyme* (1OU5) і *E1beta subunit of pyruvate dehydrogenase* (1IK6)), структуру цього фрагмента не визначено через його невпорядкованість у кристалах цих білків. Можна припустити, що штучна послідовність Met1-Met45 взагалі має невпорядковану просторову структуру і, отже, містить малу кількість регулярних елементів вторинної структури. Це також узгоджується з результатами 11-ти програм передбачення вторинної структури, які детектують лише одну α -спіраль завдовжки 10—15 а.з. в середині цієї ділянки на фоні в цілому нерегулярної структури (рис. 3). Таким чином, ділянка Met1-Met45 може мати лише незначний вплив на дані, отримані з КД-спектрів білка YCD3k, у бік підвищення вмісту α -спіральної конформації.

Модель просторової структури С-модуля, яка відповідає останнім 166 а.з. рекомбінантного білка YCD3k (Val363-Ser528 за нумерацією TyrRS бика), була раніше отримана нами на основі структури гомологічного білка ЕМАР ІІ людини за допомогою серверу Swiss-Model з наступною оптимізацією структури моделі за методом молекулярної динаміки [7]. З атомних

координат цієї моделі виведено межі елементів вторинної структури з допомогою трьох різних алгоритмів розпізнавання — DSSP, STRIDE і P-SEA. Ці програми розпізнають у структурі моделі С-модуля відповідно 58, 61 та 60 а.з. з 166 а.з., які знаходяться у конформації β -тяжів (E). Тобто всі алгоритми дають практично однакові значення, що вкладаються в діапазон 34,9—36,7 %. У спіральній конформації (α - та 3_{10} -спіралі, H+G) зазначені вище програми розпізнають відповідно 15 (9,0 %), 16 (9,6 %) та 5 а.з. (3,0 %), тобто програма P-SEA дає знижену оцінку вмісту спіралей. Більш деталізовані значення окремих елементів вторинної структури (розпізнавання за стандартним алгоритмом DSSP) становлять: β -тяжі (E) — 58 а.з. (34,9 %), α - і 3_{10} -спіралі (H+G) — 15 а.з. (9,0 %), вигини (S) — 16 а.з., повороти (T) — 25 а.з., β -вип'ячування (B) — 7 а.з., нерегулярна конформація — 43 а.з.

Порівняння результатів і обговорення. Результати, отримані з допомогою незалежних підходів (експериментальних і теоретичних), узагальнено із зазначенням даних на послідовність С-модуля завдовжки 166 а.з. (рис. 4). Якщо знехтувати даними застарілих програм передбачення GOR1 і GOR2, найкраще наближення до результатів виведення з просторової структури показали програми передбачення PHD та GOR4, які відповідно дещо занижили та перебільшили вміст β -конформації. Це узгоджується з літературними даними про більш високий прогностичний потенціал зазначених програм [11, 12]. Ці значення є проміжними між даними аналізу КД-спектрів та визначення вторинної структури лише з амінокислотної послідовності. Значення вмісту елементів вторинної структури, отримані з обчислення КД-спектрів (48 % β -тяжів і 4 % α -спіралей), дають дещо завищене передбачення β -тяжів порівняно з даними передбачення з послідовності та виведення з моделі просторової структури С-модуля TyrRS.

Слід наголосити на тому, що в межах ділянки Asp190-Gln198 усі програми визначення вторинної структури (крім DSC) прогнозують наявність α -спіралі, тоді як у просторовій структурі моделі присутній β -тяж. Це пояснюється тим, що всі ці програми враховують лише вплив безпосереднього оточення кожно-

го амінокислотного залишку і не зважають на далекодуючі просторові взаємодії між залишками, розташованими в різних місцях послідовності. Необхідно також зазначити, що програми PHD і DSC більш точно локалізують межі передбачуваних елементів.

Проведено порівняння отриманих даних з просторовою структурою гомологічного С-модуля TyrRS людини (92,9 % ідентичності), визначеною методом рентгеноструктурного аналізу (PDB код 1NTG) [8]. С-модуль людини за алгоритмами DSSP, STRIDE і P-SEA містить відповідно 38,5, 40,2 і 43,8 % β -тяжів, а спіралей ($\alpha+3_{10}$) — 8,9, 8,9 і 3,6 %.

Отже, отримані експериментальні дані за-

довільно узгоджуються з передбаченням вторинної структури С-модуля безпосередньо з амінокислотної послідовності, а також з її детектуванням у моделях просторової структури С-модуля. Це, зокрема, свідчить про нативний стан рекомбінантного білка YCD3k в отриманих препаратах та адекватність умов його зберігання.

Подяки. Роботу підтримано NATO Linkage grant HTECG.LG 974684 та NATO Computer networking supplement No. 976022 до NATO linkage grant.

УДК 577.152.6:577.322

Надійшла до редакції 30.09.2005 р.

Secondary structure analysis of cytokine-like C-terminal module of mammalian (*Bos taurus*) tyrosyl-tRNA synthetase using methods of circular dichroism spectra and computer modeling

К.А. Odynets, А.І. Kornelyuk

Institute of Molecular Biology and Genetics, NAS of Ukraine
150 Zabolotny Str., Kyiv, 03143, Ukraine

Abstract. COOH-Terminal module of mammalian cytoplasmic tyrosyl-tRNA synthetase (TyrRS) is an additional RNA-binding domain, which also has cytokine-like activities in isolated state after its proteolytic release from a catalytic core of the enzyme. In the article a comparative secondary structure analysis of this bifunctional C-module of bovine TyrRS (*Bos taurus*) is fulfilled using experimental and theoretical methods. The experimental determination of secondary structure was carried out by the method of circular dichroism (CD) spectra analysis for recombinant C-module. The theoretical prediction was carried out using: 1) the programs of protein secondary structure prediction from its amino acid sequence, and 2) the deduction algorithms of secondary structure elements from atomic coordinates (DSSP, STRIDE and P-SEA) of a model of C-module spatial structure, obtained by homology modeling technique using human cytokine EMAP II. Experimental CD data for C-module of TyrRS confirms the native state of protein YCD3k in obtained preparations.

Key words: tyrosyl-tRNA synthetase, C-terminal module, secondary structure prediction, circular dichroism spectra.

Перелік літератури

1. Корнелюк А.І. Структурное и функциональное исследование тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка. — 1998. — Т. 14, № 4. — С. 349—359.
2. 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 радиоактивно меченного гибрида зонда // Биополимеры и клетка. — 1997. — Т. 13, № 2. — С. 121—126.
3. 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. — Т. 13, № 6. — С. 474—478.
4. Дубровский А.Л., Браун Дж., Корнелюк А.И., Мюррей Дж.К., Мацука Г.Х. Бактериальная экспрессия полноразмерных и усеченных форм цитокина ЕМАР-2 и цитокин-подобного домена тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка. — 2000. — Т. 16, № 3. — С. 229—235.
5. Kornelyuk A.I., Tas M., Dubrovsky A., Murray C. Cytokine activity of the non-catalytic EMAP-2-like domain of mammalian tyrosyl-tRNA synthetase // Биополимеры и клетка. — 1999. — Т. 15, № 2. — С. 168—172.
6. Голуб А.Г., Одынец К.А., Ныпорко А.Ю., Корнелюк А.И. Моделирование пространственной структуры С-концевого цитокин-подобного модуля цитоплазматической тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка. — 2000. — Т. 16, № 6. — С. 515—524.
7. 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 // Фізика живого. — 2003. — Т. 11, № 2. — С. 61—71.
8. Yang X.-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 // *Helvetica Chimica Acta*. — 2003. — Vol. 86. — P. 1246—1257.

9. *Andrade M.A., Chacón P., Merelo J.J., Morán F.* Evaluation of secondary structure of proteins from UV circular dichroism using an unsupervised learning neural network // *Prot. Eng.* — 1993. — Vol. 6, No 4. — P. 383—390.

10. *Merelo J.J., Andrade M.A., Prieto A., Morán F.* Proteinotopic feature maps // *Neurocomputing*. — 1994. — Vol. 6, No 2. — P. 443—454.

11. *Combet C., Blanchet C., Geourjon C., Deléage G.* NPS@: Network Protein Sequence Analysis // *TIBS*. — 2000. — Vol. 25, No 3. — P. 147—150.

12. *Rost B., Sander C., Schneider R.* PHD — an automatic mail server for protein secondary structure prediction // *Comput. Appl. Biosci.* — 1994. — Vol. 10, No 1. — P. 53—60.

13. *Kim Y., Shin J., Li R., Cheong C., Kim K., Kim S.* A novel anti-tumor cytokine contains an RNA binding motif present in aminoacyl-tRNA synthetases // *J. Biol. Chem.* — 2000. — Vol. 275, No 35. — P. 27062—27068.

14. *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. — Vol. 20, No 3. — P. 570—578.

15. *Guex N., Peitsch M.C.* SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling // *Electrophoresis*. — 1997. — Vol. 18, No 15. — P. 2714—2723.

16. *Kabsch W., Sander C.* Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features // *Biopolymers*. — 1983. — Vol. 22, No 12. — P. 2577—2637.

17. *Labesse G., Colloc'h N., Pothier J., Mornon J.P.* P-SEA: a new efficient assignment of secondary structure from C alpha trace of proteins // *Comput. Appl. Biosci.* — 1997. — Vol. 13, No 3. — P. 291—295.

18. *Douquet D., Labesse G.* Easier threading through web-based comparisons and cross-validations // *Bioinformatics*. — 2001. — Vol. 17, No 8. — P. 752—753.

19. *Johnson W.C.* Analyzing protein circular dichroism spectra for accurate secondary structures // *Proteins*. — 1999. — Vol. 35, No 3. — P. 307—312.

20. *King S.M., Johnson W.C.* Assigning secondary structure from protein coordinate data // *Proteins: Structure, Function and Genetics*. — 1999. — Vol. 35, No 3. — P. 313—320.

21. *Pancoska P., Janota V., Kiederling T.A.* Novel matrix descriptor for secondary structure segments in proteins: Demonstration of prediction from circular dichroism spectra // *Analytical Biochemistry*. — 1999. — Vol. 267, No 1. — P. 72—83.

22. *Sreerama N., Venyaminov S.Y., Woody R.W.* Estimation of number of a helical and b strand segments in proteins using CD spectra // *Protein Science*. — 1999. — Vol. 8, No 2. — P. 370—380.

23. *Sreerama N., Venyaminov S.Y., Woody R.W.* Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis // *Anal. Biochem.* — 2000. — Vol. 287, No 2. — P. 243—251.