



(12)

CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: **a 2011 00377**

(22) Data de depozit: **20.04.2011**

(66) Prioritate internă:

05.11.2010 RO a 2010 01055

(41) Data publicării cererii:

28.10.2011 BOPI nr. **10/2011**

(71) Solicitant:

• **INSTITUTUL NAȚIONAL DE
CERCETARE-DEZVOLTARE PENTRU
ȘTIINȚE BIOLOGICE,
SPLAIUL INDEPENDENȚEI NR.296,
SECTOR 6, BUCUREȘTI, B, RO**

(72) Inventatori:

• **BUTU ALINA, STR. DEALUL ȚUGULEA
NR.32-36, BL.15, SC.A, AP.12, SECTOR 6,
BUCUREȘTI, B, RO;**
• **BUTU MARIAN, STR. DEALUL ȚUGULEA
NR.32-36, BL.15, SC.A, AP.12, SECTOR 6,
BUCUREȘTI, B, RO**

(54) **METODĂ DE ANALIZĂ A CONFORMAȚIEI SECVENȚELOR
PROTEICE**

(57) Rezumat:

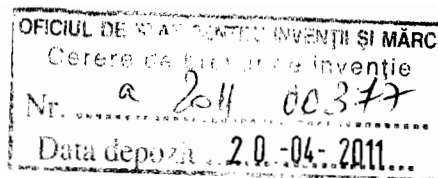
Prezenta invenție se referă la o metodă de analiză a conformației secvențelor proteice, utilă în investigarea structurii și funcțiilor proteinelor. Metoda conform invenției determină modificările conformaționale ale secvenței proteice, care apar de-a lungul simulării dinamicii

moleculare, prin calculul variațiilor numărului total de atomi aflați într-un spațiu cu dimensiuni predefinite.

Revendicări: 1

Figuri: 3





Descriere

Metoda descrisa in aceasta inventie este dedicata analizei modificarilor conformationale aparute in dinamica moleculara a proteinelor.

Proteinele sunt constituentii ai organismelor vii cu cel mai inalt grad de complexitate si de varietate moleculara. Ele prezinta specificitate de organ si de specie. Denumirea lor deriva de la cuvantul grecesc *proteios* care inseamna "cel dintai, primul". Datorita rolurilor fundamentale pe care le indeplinesc in organismele vii, proteinele se bucura de o atentie deosebita din partea cercetatorilor.

Funcțiile îndeplinite de proteine acopera intreg spectrul de activitate ale celulelor. Astfel, proteinele:

- au rol structural major, ele constituind materialul din care sunt constituite toate structurile celulare, membrane, organe celulare, precum si materialul intercelular.
 - exercita actiuni catalitice, determinand varietatea nesfarsita de reactii biochimice si specificul transformarilor chimice din organismele vii.
 - contractile sunt instrumente cu ajutorul carora organismele vii indeplinesc activitatea contactila si de locomotie.
 - indeplinesc functii reglatoare, ele pot stoca si transmite mesaje chimice.
 - indeplinesc functii de transport si de depozitare a unor compusi chimici ca ioni metalici, vitamine, oxigen, dioxid de carbon.
 - au sarcini de a apara organismul impotriva unor corpi straini, macromolecule, virusuri, bacterii.
- Reactiile imunologice sunt mediate de o clasa de proteine specializate.

Pentru a înțelege metoda de analiza descrisa în aceasta inventie trebuie explicat ce înseamnă o proteină. Există două modalități de a descrie o proteină: "de la complex la simplu" sau "de la simplu la complex". Prima metodă poate fi comparată cu a privi un obiect la microscop și a crește gradat factorul de mărire. Prin această metodă se pornește de la întreaga proteină și complexitatea structurii descrește prin intrare în detaliu. A doua metodă începe cu nivelul de bază al atomilor individuali și pierde detaliile odată cu descreșterea măririi. Aici se folosește abordarea "de la simplu la complex".

Din punctul de vedere al fizicianului, proteina este un set de sfere moi. Aceste sfere pot fi bleu, albe, albastre, roșii, galbene etc, culoarea diferind în funcție de masă și rază. Acestea sunt conectate cu una până la patru alte sfere prin legături rigide scurte. Avem o descriere unică a proteinei atunci când se cunoaște culoarea sferelor și modul în care sunt conectate cu alte sfere (topologia). De fapt, descrierea anterioară se aplică oricărei molecule. Proteinele sunt un subset al tuturor moleculelor posibile, deoarece nu se întâlnesc toate combinațiile de sfere în proteine. În lumea reală, sferele reprezintă atomi și legăturile sunt legături chimice (covalente).

Din punctul de vedere al chimistului o proteină poate fi comparată cu un șir de mărgelile. Există 20 de tipuri de mărgelile, fiecare cu formă și mărime diferită. Forma unei mărgelile nu este fixă, ea poate varia, dar numărul de forme este limitat. Mărgelile sunt conectate cu una sau două alte mărgelile (unele mărgelile pot realiza o a treia conexiune). Uzual mărgelile sunt numite reziduuri.

Toate reziduurile sunt levogire și au formula generală prezentată în Fig 1. 20 -04- 2011

Fiecare reziduu conține două părți, lanțul principal care poate să se conecteze prin legături chimice covalente de unul sau două reziduuri și un lanț lateral, notat cu litera R în Fig. 1. Atomul central carbon de care se leagă grupul lateral se numește carbon alfa ($C\alpha$). Aminoacizii se leagă unul de altul printr-o legătură peptidică formată între azotul teminal și carbonul terminal al aminoacidului adiacent cu eliminarea unei legături de apă și formează un lanț polipeptidic (Fig. 2).

O moleculă de proteină poate fi alcătuită din câțiva aminoacizi până la câteva mii de aminoacizi legați în acest fel. După formarea proteinei rămân două grupări nelegate și sunt cunoscute ca N și C terminal. Carbonul $C\alpha$ și atomii C și N adiacenți ai fiecărui reziduu aminoacid definesc lanțul principal al proteinei.

Lanțul principal este identic în toate reziduurile, iar lanțul lateral diferențiază reziduurile. Datorită alcătuirii lor diferite, reziduurile au diferite proprietăți. Unele reziduuri atrag apa (reziduuri hidrofile sau polare), altele resping apa (reziduuri hidrofobe). Reziduurile hidrofile sunt încărcate electric având sarcina $\pm 1 e$.

Din punctul de vedere al biologului proteina este o moleculă complexă care realizează o anume funcție. Domeniul funcțiilor proteinelor este vast, unele au rol structural, cum este keratina, altele se găsesc în peretele celular și facilitează transportul particulelor mici cum sunt ionii sau apa (porinele), altele catalizează reacții (enzimele), iar alte proteine reglează funcția altor proteine. Biologii încearcă să înțeleagă relațiile dintre proteine, dintre proteine și alte părți ale celulei și cum anumite procese sunt reglate și organizate într-o descriere spațio-temporală.

O moleculă proteică este un lanț liniar de aminoacizi care se pliază într-o formă reproductibilă.

Proteina poate fi văzută la diferite nivele de detaliere. Anterior, proteina a fost descrisă din punct de vedere atomic și din punct de vedere al reziduurilor. O alta descriere este funcție de grupurile de reziduuri care formează structura locală, numită elemente de structură secundară. Structura secundară descrie conformația lanțului principal al proteinei. Cel mai important element de structură este α -helix .

Reprezentarea tuturor atomilor este folosită în general numai pentru peptide mici, care nu au o structură secundară bine definită. În termenii structurii secundare sunt descrise proteinele.

Proteinele pot fi împărțite în două categorii după formă și funcție: globulare și fibroase. Proteinele globulare participă, în general, la procesele chimice dintr-un organism, de exemplu enzimele și proteinele redox, în care proteinele fibroase joacă un rol mai mult mecanic și structural, de ex. actina și miozina din mușchi și collagenul din tendoane. Cea din urmă categorie tinde să fie foarte insolubilă și cristalizarea este aproape imposibilă. Pe de parte proteinele globulare sunt mult mai bune candidate pentru cristalizare datorită solubilității lor și sunt subiectul majorității studiilor cristalografice macromoleculare și a celor de simulare a dinamicii moleculare.

Pentru investigarea structurii și funcțiilor proteinelor se încearcă dezvoltarea de noi tehnici și metode de analiză a dinamicii moleculare a acestora. Astăzi este o tendință generală în domeniul științelor biologice să se interpreteze datele experimentale într-un mod din ce în ce mai complex. Acest fapt implică o creștere a acurateții calculului, din ce în ce mai elaborate și mai complexe.

Un punct important al dinamicii este faptul că se urmărește evoluția în timp a sistemului și se prezintă complet informația despre dinamica sistemului. Datele experimentale sunt de obicei valabile într-un spațiu vector-frecvență. Informația obținută prin dinamica moleculară în timp și spațiu real va fi deci transformată Fourier în raport atât cu spațiul cât și cu timpul pentru o comparare utilă cu rezultatele experimentale. Pentru a realiza analiza dinamicii, se scriu în programul de simulare instrucțiuni de descărcare periodică a pozițiilor și vitezelor tuturor particulelor pe un mediu de stocare. Aceste date sunt analizate mai târziu cu alte programe pentru extragerea mărimilor de interes. În acest moment sunt utilizate pentru determinare modificărilor conformationale mai multe metode: prin evaluarea evoluției unghiurilor diedre, evaluarea distanțelor C α -C α și altele. Toate aceste metode sunt complementare între ele ca și metoda propusă de noi. În prezenta invenție se propune analizarea traiectoriilor prin calcularea variației numărului de atomi ce se găsesc într-un spațiu cu dimensiuni predefinite. Variația numărului total de atomi va indica dacă există o modificare a conformației proteinei în spațiul predefinit și momentul în care are loc această modificare.

În experimentele de simulare a dinamicii moleculare a unei secvențe proteice sunt înregistrate cronologic într-un fișier pozițiile tuturor atomilor proteinei. Astfel, la momentul initial, notat cu t_0 , avem setul de poziții $(p_{0,1}, p_{0,2}, \dots, p_{0,n})$ unde n este egal cu numărul total de atomi ai proteinei; la momentul t_1 , avem setul de poziții $(p_{1,1}, p_{1,2}, \dots, p_{1,n})$... Generalizând, la momentul oarecare t_m al simulării pozițiile atomilor din secvența proteică pot fi notate astfel $(p_{m,1}, p_{m,2}, \dots, p_{m,n})$.

Aplicăm metoda de analiză prezentată aici pentru un atom de referință notat cu a , unde a este un număr real, întreg, cuprins între 1 și n . Considerăm un volum definit V centrat pe atomul a și, folosind seturile de poziții din fișierul rezultat în urma simulării, calculăm pentru fiecare moment m de timp valorile $p_{m,ax} = p_{m,a} - p_{m,x}$, unde x ia valori de la 1 la n și aplicăm condiția dacă $|p_{m,ax}| \subset V$, $ax=1$, dacă $|p_{m,ax}| \not\subset V$, $ax=0$. Calculăm pentru fiecare moment de timp m numărul de atomi N_m care se găsesc în volumul V , folosind formula $N_m = \sum_{x=1}^n ax$. Variațiile funcției $N_m(m)$ corespund modificărilor conformationale aparute în volumul definit în jurul atomului a . Astfel, în reprezentarea grafică a funcției $N_m(m)$, se identifică picurile (varfurile) curbei și se determină cu exactitate momentele de timp la care apar modificările conformationale.

Mentionăm că metoda se poate aplica și pentru volume definite de diverse forme, exemplu: cub, sferă, dodecahedron, etc. Indiferent de forma volumului definit în jurul atomului de referință, se păstrează condiția de contorizare doar a atomilor ce se găsesc în acest volum la momentul de timp m .

Un exemplu de grafic rezultat este prezentat în fig. 3 în care sunt încercuite cu roșu zonele în care apar modificări conformationale.

Revendicare

Metoda de analiza a conformatiei secventelor proteice este **caracterizata prin aceea ca** determina modificarile conformationale ale secventei proteice aparute de-a lungul simularii dinamicii moleculare prin calcularea variatiilor numarului total de atomi aflati intr-un spatiu cu dimensiuni predefinite.

20-04-2011

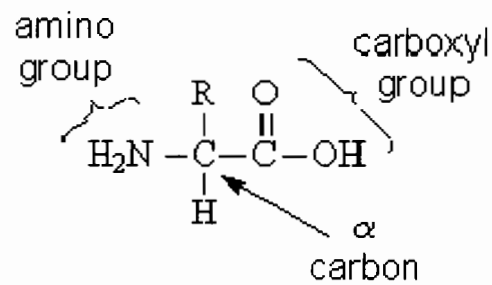


Fig. 1. Formula generală a aminoacizilor



Fig. 2. Reprezentarea lanțului principal (verde) și a lanțului lateral (gri) într-o peptidă

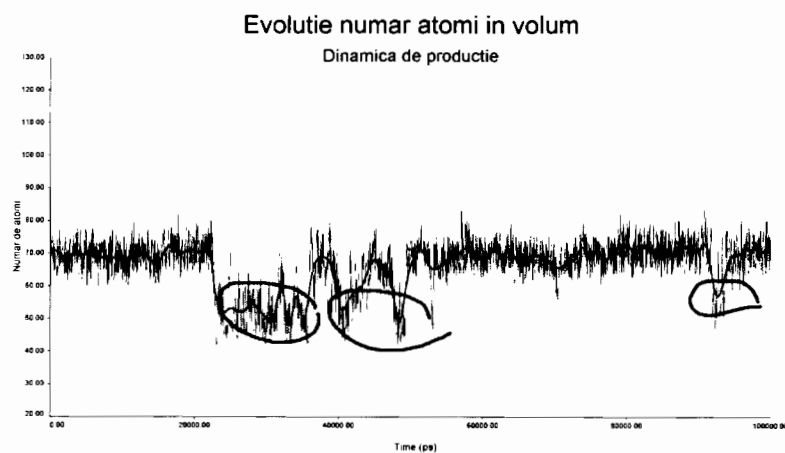


Fig. 3. Reprezentarea variației numărului de atomi dintr-un volum predefinit