

**INTRODUCERE ÎN MODELELE MATEMATICE  
ȘI SIMULĂRILE CINETICE A PROCESELOR BIOLOGICE  
(Revista literaturii)**

**Viorel Munteanu**

Laboratorul Inginerie Tisulară și Culturi Celulare, USMF "Nicolae Testemițanu"

**Summary**

***An introduction to mathematical models and kinetic simulations of biological processes***

The function of living cells is controlled by complex regulatory networks that are built of a wide diversity of interacting molecular components. Recent technological developments have enabled single-cell measurements of cellular macromolecules which can shed new light on processes underlying gene expression. Correspondingly, there is a need for the development of theoretical tools to quantitatively understand gene expression and its consequences for cellular processes. The review focuses attention on the necessity of modeling and mentions the basic stages in the mathematical dynamical modeling in biological processes.

**Introducere**

Modelele matematice dinamice a rețelelor biochimice și genetice reprezintă abstracții a realității, fiind construite în special pentru a ne ajuta a înțelege. Toate modelele sunt construite cu accent spre înțelegerea anumitor aspecte a obiectelor de studii, celelalte fiind abstractizate.

Cercetătorii regulat utilizează modele reale pentru cercetări, acestea pot fi mai simple sau mai complexe, precum organismele model sau modelele patologice. Cercetătorii la fel utilizează modelele conceptuale, acestea reprezentând descrieri verbale a sistemelor și a operării acestora, deseori fiind comunicate prin diagrame care ilustrează un set de componenți și calea prin care aceștia interacționează. Aceste diagrame a interacțiunilor sau modelele „proiectate” joacă rolul central în reprezentarea cunoștințelor proceselor celulare. Dezavantajul acestor modele „proiectate” este că acestea pot fi lipsite de caracteristici importante privind operarea și evoluția sistemului, în special când rețeaua de interacțiuni implică feedback-urile. Utilizarea descrierilor matematice a sistemelor permite eliminarea incertitudinilor din operarea modelului, precum și permite reprezentarea cantitativă a fiecărui component din sistem. Astfel, cunoscând datele cantitative, diagrama de interacțiuni poate fi folosită pentru a formula modelul matematic dinamic.

Descrierile cantitative a interacțiunilor moleculare de obicei apelează la legile fizicii și chimiei. Modelele rezultate sunt *mecanistice* – acestea descriu mecanismele care guvernează operarea sistemului. Modelele mecanistice pot fi contrastate cu modelele *statistice* care sunt îndreptate spre caracterizarea unui set de date (denumite de obicei și modele descriptive). Fiecare componentă a modelului mecanistic reprezintă unele aspecte a sistemului care urmează a fi studiate, modificările în operarea modelului, astfel, mimică modificările în sistemul real.

Investigarea modelelor matematice urmează prin două părți complementare. Metoda cea mai directă este simularea modelului, în care modelul este utilizat pentru a prezice comportamentul sistemului, în anumite condiții. Simulările se referă la experiențele *in silico*, acestea se efectuează pe baza calculatorului pentru a explora comportamentul sistemelor biologice. Aceste simulări sunt executate de diverse pachete de software, care permit efectuarea diferitor analize a proprietăților sistemului.

O alternativă a simulărilor este investigarea directă a modelelor care oferă informații și cunoștințe privind potențialul comportament a acestora. Această metodă de analiză a modelelor implică în general tehnici matematice sofisticate.

Aceste tehnici permit înțelegerea comportamentului (existenței) sistemului, ceea ce nu poate fi obținut prin simulările simple. Dacă simulările arată *cum* se comportă un sistem, analiza

modelului ne oferă informații privind *de ce* sistemul se comportă în modul dat. Aceste analize pot să releveze conexiunile ne-intuitive dintre structurile sistemului și comportamentul acestora.

În acest articol ne propunem să aducem o introducere generală privind sistemele biologice și simulările acestora, necesitatea simulărilor și etapele de bază în planificarea unui studiu de simulare.

### **1. Necesitatea modelelor matematice dinamice**

Precum s-a menționat, digramele de interacțiuni, tipic, omit ambiguități cu respect la comportamentul sistemului, mai ales când sunt implicate mecanismele feedback. Cu atât mai mult, odată cu creșterea elementelor ce interacționează în rețea, devine foarte dificil de a avea și menține o înțelegere intuitivă a comportamentului global a sistemului. Aceasta este o provocare a biologiei sistemelor și este concisă prin expunerea ”celulele sunt sisteme complexe”.

Termenul de sistem are sensul de ceva context-dependent, însă tipic se referă la o colecție de componente ce interacționează. Termenul de sistem presupune ”orice ce-și vorbește” (Kevin Kelly’s în *Out of Control*).

Printre numeroasele componente ce interacționează, o altă trăsătură definitorie a sistemului este *hotarul* (marginea). Un sistem constă dintr-un set de componente; oricare alt element ce nu face parte din acest set de componente, nu este parte a sistemului și reprezintă o parte a „mediului extern”. În cazul rețelelor de reglare genetice, sistemul este definit exclusiv de elemente ce interacționează între ele și comunică în același timp cu „mediul extern” constituind sisteme de intrare-ieșire.

Sistemele sunt mai complexe, odată ce comportamentul global a acestora nu pot fi intuitiv înțeles sau prezis, în termeni de interacțiune sau componente individuale. O trăsătură definitorie a sistemelor complexe este că natura calitativă a comportamentului acestora poate să depindă de diferențele cantitative în structura modelului. Două trăsături esențiale a sistemelor complexe sunt interacțiunile ne-liniare și buclele feedback (negative și pozitive).

### **2. Utilizarea modelelor matematice dinamice**

Modelele matematice dinamice servesc ca ajutor în numeroase tipuri de investigații biologice. Procesul de proiectare și construire a modelului necesită considerații critice a mecanismului care stă la baza procesului biologic. Aceasta echivalează cu un proces reflectiv riguros, care poate să conducă la evidențierea unor contradicții în modelul proiectat și evidențierea unor lacune în cunoștințele existente. Odată ce modelul a fost construit, acesta servește drept o descriere transparentă a sistemului și poate fi fără echivoc comunicat. Astfel, un model trebuie să recapituleze comportamentul unui sistem și să însușeze concis cunoștințele acumulate privind acesta, în plus modelele trebuie să poată să fie expuse reproducerii.

Atât modelul proiectat cât și modelul matematic reprezintă manifestări a ipotezelor: acestea corespund descrierilor posibile a sistemelor și operărilor acestora. Avantajul modelului matematic este că acesta este o „ipoteză în acțiune”, în sensul că starea dinamică a acestuia poate fi nemijlocit investigată. Simulările modelului nu pot înlocui experiențele de laborator, însă modelul permite testarea operării sistemului, ceea ce cel mai des nu este posibil de efectuat în condiții de laborator. Simulările pot fi efectuate rapid (cel mai des în secunde sau minute) și nu suportă costuri reale. Comportamentul sistemului poate fi explorat în condiții care niciodată nu ar fi putut obținute în laborator și fiecare aspect a comportamentului sistemului poate fi observat în orice moment de timp. În plus, analiza modelului permite observările dintre calea de evoluție a sistemului și motivele trecerii prin aceste stări, astfel asigurând o legătură dintre structura rețelei și comportamentul sistemului.

Deoarece modelul reprezintă o ipoteză, rezultatele investigațiilor modelului reprezintă la fel ipoteze. Simulările nu pot definitiv prezice comportamentul celulelor, însă acestea pot servi drept un ghid prețios pentru planificarea experiențelor, indicând căile promițătoare pentru

investigații, sau indicînd asupra contradicțiilor dintre înțelegerea noastră a sistemului (încorporat în model) și observațiile din laborator. De fapt, identificarea acestor contradicții reprezintă beneficiul prioritar a modelării: deoarece un model poate fi exhaustiv investigat, aceasta rezultă în faptul că rezultatul negativ – incapacitatea modelului de a reproduce observațiile experimentale – pot fi considerate ca falsificări a ipotezelor pe care a fost construit modelul. Astfel ducînd la o revedere (revizuire), analiză și rafinare a ipotezelor biologice și consecutiv o alegere corectă a modelului, care poate fi apoi testat în contextul experiențelor adiționale. Acest proces iterativ conduce la o îmbunătățire continuă a înțelegerii sistemului, proces numit „ciclu virtual”. Scopul final al eforturilor de modelare este descrierea predictivă a sistemului; simulările constituind reprezentări precise a comportamentului real.

În biologia sintetică o aplicare tot mai largă o are proiectarea bazată pe modele (*model-based design*). Deși modelele rețelelor celulare posedă o putere de prezicere limitată, acestea sunt de mare folos și preț pentru ghidarea alegerii componentelor și sugerarea experiențelor celor mai efective pentru testarea performanței sistemului.

### 3. Modelarea cinetică a sistemelor biologice

Reacțiile care au loc între un grup de reactanți definesc rețeaua reacțiilor cinetice. Tradițional, setul de reacții este descris de ecuații ordinare neliniare sau ecuații diferențiale parțiale, soluțiile cărora reprezintă dinamica acestor procese. Mărimea acestor rețele poate să varieze considerabil, în unele cazuri, sute de mii de reacții sunt necesare pentru a descrie complexitatea fenomenelor cuplate.

Rețelele biologice deseori prezintă bucle feedback, bifurcări și oscilații, fenomene care conduc la stări dinamice foarte variate și complexe. Deși prezintă numeroase asemănări, două aspecte în particular deosebesc rețelele biologice de rețelele întâlnite în alte domenii. În primul rînd, în sistemele biologice, numărul (cantitatea) unor specii moleculare este foarte mic, în al doilea rînd, sistemele biologice tind să aibă structuri ierarhizate bine organizate.

Calea tradițională de modelare a evoluției în timp a populațiilor moleculare în sistemul de reacții este utilizarea ecuațiilor diferențiale ordinare (EDO) de ordinul unu, numite și ecuații de viteză a reacțiilor (EVR). EVR sunt de obicei înferate din argumente fenomenologice cu presupunerea (ipoteza) că sistemul este bine omogenizat sau spațial omogen. Aceste ecuații sunt neliniare în prezența reacțiilor bimoleculare, ceea ce este aproape întotdeauna. Din cauza ratelor foarte diferite, EVR pot manifesta diferite scări de timp, condiție cunoscută ca rigiditate dinamică. Uneori, în cazul unor presupuneri, cu scopul simplificărilor, precum stările de echilibru sau echilibrul parțial, EVR pot fi reduse la un sistem de ecuații diferențiale algebrice (EDA).

Utilizarea metodei deterministe poate fi insuficientă pentru sistemele celulare, cînd problema de bază este de natură stohastică. Fluctuațiile concentrațiilor moleculelor pot fi destul de mari și acesta necesită o tratare specială a rețelelor biologice [1-4]. Acestea sunt adevărate și se referă în special la speciile moleculare implicate în reglarea transcripțională a expresiei genelor. În unele astfel de cazuri, la atingerea unei limite (număr mic de copii de molecule), modelele bazate pe ecuațiile diferențiale nu pot fi utilizate, întrucît presupunerea concentrațiilor moleculare în dinamică nu mai este validă. În aceste cazuri o abordare mai generală, bazată pe metodele stohastice, în care fluctuațiile naturale sunt tratate în direct într-un mod explicit este susținută de analizele cantitative a rețelelor biologice [5-7]. Aceste cercetări au arătat că abordările stohastice discrete sunt, cel mai des, mai potrivite pentru studierea sistemelor biologice. Deși simulările stohastice discrete oferă un cadru de modelare mai realist, acestea sunt de obicei foarte lente și costisitoare în resurse informatice. În baza primelor lucrări din acest domeniu [1, 2, 8] au fost dezvoltate noi metode matematice și algoritme numerice pentru a depăși unele probleme de calcul [5, 6, 9-23].

În linii generale, simularea cinetică a sistemelor biologice implică următoarele etape:

- i. **Identificarea problemei cinetice.** Această etapă implică determinarea variabilelor de intrare și de ieșire, precum și a celor intermediare. Sunt determinate și tabelate speciile de bază a moleculelor implicate.
- ii. **Formularea modelului.** Cele mai multe studii utilizează EVR pentru modelarea sistemelor biochimice. Aceste ecuații pur și simplu definesc schimbările în concentrații (sau echivalent în numărul de molecule) în funcție de timp și locație.
- iii. **Alegerea unei metode.** La această etapă se decide asupra alegerii formulării deterministe sau stohastice. Mai rar se alege o abordare hibridă dintre aceste modele.
- iv. **Simularea.** Integrarea deterministă a EVR implică alegerea unui algoritm adecvat dintr-o gamă largă de metode numerice pentru EDO și EDA. Pentru acesta, sunt disponibile o gamă largă de softuri înalt eficiente și de încredere.
- v. **Analizele traiectoriei.** Cursul timpului obținut din simulări este catalogat și combinat pentru a deriva distribuția statistică a cantităților de ieșire și intrare, monitorizarea concentrațiilor speciilor și fluxul maselor în sistem sau probabilitatea distribuției reacțiilor. Se acordă atenție la studiul sensibilității la variațiile parametrilor.

Comportamentul primar a modelului matematic dinamic corespunde speciilor moleculare în diagrama de interacțiuni corespunzătoare. Abundența fiecărei specii este atribuită la starea variabilei din model. Colecția tuturor acestor stări a variabilelor reprezintă starea globală a sistemului. Aceasta oferă o descriere completă a condițiilor sistemului la orice moment de timp, astfel, comportamentul dinamic a modelului reprezintă o colecție de-a cursul timpului a stărilor variabilelor.

În afară de stările variabilelor, modelele includ și parametri. Parametrii modelului caracterizează interacțiunile dintre componentele sistemului precum și dintre acestea cu mediul. Comportamentul stărilor variabilelor este astfel dependent de valorile parametrilor modelelor. Exemple de parametri a modelelor includ constantele de asociere, ratele de expresie maximală, ratele de degradare și concentrațiile moleculare tampon. O modificare în valoarea parametrului modelului corespunde unei schimbări în condițiile mediului sau în interiorul sistemului. În consecință, parametrii modelului sunt de obicei considerate valori constante pe parcursul simulărilor, acestea putând varia numai pentru extrapolarea și studiul comportamentului sistemului în condiții de perturbații sau condiții externe alterate (ex. condiții experimentale diferite).

Simulările modelelor dinamice ne arată în ce mod sistemul se modifică pe parcursul timpului. Modelele proceselor biologice, aproape întotdeauna ajung pînă la urmă în stările de echilibru. Cel mai frecvent, modelele prezintă o stare operatorie persistentă, denumită stare de echilibru, la unele sisteme, acesta fiind reprezentată de oscilații. Perioada de timp care conduce de la starea inițială la această stare îndelungată (starea asimptotică), se referă la perioada tranzientă. Această perioadă merită o atenție deosebită, deoarece aceasta reflectă răspunsul imediat a sistemului la perturbații. Merită de atras atenție și la stările de echilibru a sistemului, deoarece acestea reflectă condițiile predominante a sistemului peste perioade semnificative de timp.

O relație dintre componentele sistemului este numită liniară dacă reprezintă o simplă proporționalitate. De exemplu, variabila  $x$  și  $y$  sunt în relație liniară determinată de ecuația  $x=ky$ , unde  $k$  este o constantă fixă. În această relație o dublare a lui  $x$  va conduce la o dublare a lui  $y$ , indiferent de valoarea acestuia, ceea ce indică asupra unui grad de extrapolare a stărilor acestora. Relațiile liniare care implică mai mult decît două variabile sunt la fel transparente, de ex.  $x = k_1y + k_2z$ . Un model matematic dinamic este numit liniar, dacă toate interacțiunile dintre componentele acestuia sunt în relație liniară. Aceasta este o condiție puternic restrictivă, în consecință modelele liniare prezintă o gamă îngustă de stări.

Orice relație care nu face parte din cele liniare, se referă la cele neliniare. Relațiile neliniare nu urmează un patern specific, prin urmare sunt imprevizibile și dificil asociate cu o generalitate.

Nelinearitatea care apare cel mai des în interacțiunile biochimice și genetice se referă la saturație, în care o variabilă se găsește în creștere împreună cu alta, cu o rată în descreștere, astfel, variabila dependentă tinde spre o valoare limitatoare sau asimptotică. Două tipuri de saturație sunt pe larg răspândite în rețelele genetice și cele metabolice, saturația hiperbolică, în care odată cu valoarea lui  $x$  crește și  $y$ , însă cu o rată în descreștere. Valoarea lui  $y$  astfel apropiindu-se de valoarea de limită sau asimptotică. În saturațiile sigmoide,  $y$  mai întâi crește foarte încet, împreună cu  $x$ , apoi trece printr-o perioadă de creștere rapidă, înainte de saturație după care rata creșterii scade.

Modelele matematice dinamice neliniare prezintă de obicei o gamă largă de stări. În cele mai multe cazuri, o analiză detaliată a comportamentului în ansamblu, global, a acestor modele este de obicei dificilă și foarte imensă. În schimb, ne putem adresa atenția spre aspecte specifice a stării sistemului, în particular prin limitarea atenției spre starea și comportamentul domeniilor de operare particulare, în același timp ne putem folosi de avantajul faptului că neliniaritățile pot fi approximate prin relații liniare pe fragmente mici. Această aproximare locală ne permite aplicarea instrumentelor de analiză locală pe aceste domenii limitate. Stările globale a sistemului deseori sunt constrânse și determinate de un număr mic de puncte operatorii; analiza globală a acestor puncte ne oferă o viziune și înțelegere în comportamentul global a sistemului. Aproximările locale sunt în particular utilizate în modelările biologice deoarece sistemele auto-reglabile (de ex. cele homeostatice) petrec cel mai mult timp operator în jurul acestor condiții nominale specifice.

La moment, dintre modelele existente, cele mai utilizate în procesele de simulări dinamice sunt modelele deterministe și cele stohastice. Noțiunea de determinism deduce reproductibilitatea operațiunii (comportamentului) și reprezintă o fundație pentru numeroase investigații științifice. Un model matematic este determinist dacă comportamentul acestuia este exact-reproductibil. Deși comportamentul de operare a modelului determinist este dependent de un set specific de condiții, alte forțe nu au nici o influență asupra acestuia, astfel, repetițiile simulărilor în aceleași condiții sunt întotdeauna într-un acord perfect (replicări perfecte). În contrast, modelele stohastice conduc la fenomenele aleatorii în comportamentul acestora. Comportamentul modelelor stohastice sunt determinate atât de condițiile specificate cât și de forțele imprevizibile.

#### **4. Probleme de viitor a biologiei sistemelor**

Deși, pînă în prezent s-au obținut realizări importante în domeniul biologiei sistemelor, numeroase aspecte a modelărilor rămîn a fi studiate. Acestea se referă în primul rînd la eficacitatea și efectul presupunerilor (ipotezelor) la trecerea în activitățile practice, de ex. celulele vii. La fel rămîn a fi găsite metode eficiente de studiu a rețelelor care implică intersecția mai multor bucle feedback, precum și interpretarea rolului acestora în aceste sisteme. În fine rămîne de elucidat rolul și importanța spațiului fizic în modelarea comportamentului și evoluției rețelelor biologice.

#### **Bibliografie**

1. Gillespie DT (1976) A general method for numerically simulating stochastic time evolution of coupled chemical-reactions. *Journal of Computational Physics* 22, 403–434.
2. Gillespie DT (1977) Exact stochastic simulation of coupled chemical-reactions. *Journal of Physical Chemistry* 81, 2340–2361.
3. Gillespie DT (1977) Concerning validity of stochastic approach to chemical-kinetics. *Journal of Statistical Physics* 16, 311–318.
4. Gillespie DT (1992) A rigorous derivation of the chemical master equation. *Physica A* 188, 404–425.

5. Arkin A, Ross J, and McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149 , 1633–1648.
6. McAdams HH and Arkin A (1997) Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences, USA* 94, 814–819.
7. McAdams HH and Arkin A (1999) It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics* 15, 65–69.
8. Bortz AB, Kalos MH, and Lebowitz JL (1975) New algorithm for Monte-Carlo simulation of Ising spin systems. *Journal of Computational Physics* 17, 10–18.
9. Endy D and Brent R (2001) Modelling cellular behaviour. *Nature* 409 , 391–395.
10. Gibson MA and Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A* 104 , 1876–1889.
11. Goss PJE and Peccoud J (1998) Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. *Proceedings of the National Academy of Sciences, USA* 95, 6750–6755.
12. Kastner J, Solomon J, and Fraser S (2002) Modeling a Hox gene network in silico using a stochastic simulation algorithm. *Developmental Biology* 246 , 122–131.
13. Kepler TB and Elston TC (2001) Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophysical Journal* 81, 3116–3136.
14. Rao CV and Arkin AP (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *Journal of Chemical Physics* 118, 4999–5010.
15. Simpson ML, Cox CD, and Saylor GS (2003) Frequency domain analysis of noise in autoregulated gene circuits. *Proceedings of the National Academy of Sciences, USA* 100, 4551–4556.
16. Smolen P, Baxter DA, and Byrne JH (1999) Effects of macromolecular transport and stochastic fluctuations on dynamics of genetic regulatory systems. *American Journal of Physiology-Cell Physiology* 277 , C777–C790.
17. Levchenko A (2003) Dynamical and integrative cell signaling: challenges for the new biology. *Biotechnol Bioeng* 84, 773–782.
18. Haseltine EL and Rawlings JB (2002) Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *Journal of Chemical Physics* 117, 6959–6969.
19. Haseltine EL and Rawlings JB (2005) On the origins of approximations for stochastic chemical kinetics. *Journal of Chemical Physics* 123 , 164115.
20. Puchalka J and Kierzek AM (2004) Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophysical Journal* 86, 1357–1372.
21. Fricke T and Wendt D (1995) The Markov Automaton – A new algorithm for simulating the time-evolution of large stochastic dynamic-systems. *International Journal of Modern Physics C-Physics and Computers* 6, 277–306.
22. Elf J, Donic A, and Ehrenberg M (2003) Mesoscopic reaction-diffusion in intracellular signaling. In: SPIE's *First International Symposium on Fluctuations and Noise*, pp. 114–124.
23. Elf J and Ehrenberg M (2003) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Research* 13, 2475–2484.