

## ANALIZA COMPARATIVĂ A PROTEINELOR CODIFICATE DE GENELE SISTEMULUI *ASC-Rf*

**Andrei MIDONI**

*Catedra Biologie Vegetală*

This work focussed on a comparative analysis of the codified proteins of the *CMS-Rf* system, having as purpose to find certain correlations between them and some specific properties of polypeptides.

Începând cu anul 1953, odată cu descoperirea structurii ADN-ului [7] biologia moleculară a cunoscut o dezvoltare exponențială. Apariția unui volum mare de cunoștințe noi impune, pe de o parte, necesitatea structurării acesteia în subdomenii de cercetare, iar, pe de alta – dictează utilizarea unor tehnologii de studiu și analiză mai performante. Această dilemă informațională este rezolvată cu ajutorul bazelor de date (BD), care reprezintă „adevărate depozite de informații” ce atrag anual un număr mare de cercetători.

Eficiența unei BD este determinată de cantitatea informației stocate și de prezența programelor informaționale ce tind să faciliteze atât accesul, cât și analiza datelor. În cadrul biologiei moleculare cele mai utilizate BD sunt: **The National Center for Biotechnology Information (NCBI)** [15], **DNA Data Bank of Japan (DDBJ)** [10], **DNA Data Bank** [16], **Gene Bank Database** [14], **Expert Protein Analysis System (expasy)** [12] etc. Actualmente, **NCBI**, care a fost fondată în 1988 ca o subdiviziune a bibliotecii naționale de medicină din Statele Unite ale Americii, reprezintă cel mai mare rezervor de informații cu aplicații largi în medicină, genetică și farmaceutică.

Utilizând această bază de date, am recurs la cercetarea proteinelor specifice pentru genotipurile cu androsterilitate citoplasmatică și a celor restauratoare de fertilitate la diferite specii de plante, în scopul caracterizării și identificării anumitor particularități specifice.

### Material și metode

Toate proteinele au fost selectate din baza de date NCBI [15]. Analiza unor proprietăți specifice ale proteinelor (masa moleculară, punctul izoelectric, hidrofobicitatea, sarcina electrică, frecvența aminoacizilor și structura secundară) a fost realizată prin intermediul programelor computerizate: Protein Workbench 5.0.1. [9]; Swiss Pdb Viewer ver. 4.0.1. [13]; Swissmodel.expasy.org [17]. Stabilirea tipurilor de domenii proteice (D.P.) și clasificarea acestora în noncitoplasmatică, transmembranară și citoplasmatică a fost efectuată cu ajutorul programei Phobius [11].

### Rezultate și discuții

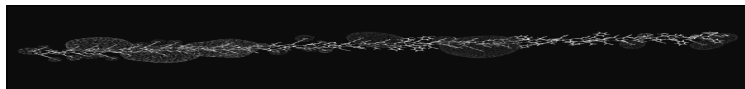
Rezultatele obținute la analiza *masei moleculare* a relevat că la toate proteinele *ASC* aceasta are o valoare mică, cu limitele inferioare de 7,12 kDa, 8,63 kDa la (*ORF67* și *ORF60* de la *Brassica oleracea*), iar cele superioare de 26,66 kDa, 43,8 kDa (*orfB-F<sub>2</sub>*, *Daucus carottai*, proteina *ASC* de la *Petunia axillaris*). La *Helianthus annuus* produsul de expresie a genei *orfH522* are o masă moleculară apropiată de valoarea medie a proteinelor *ASC*: 19,5 kDa. La proteinele restauratoare de fertilitate a fost observat faptul că acestea au mase moleculare mari, cu o valoare medie de 69,58 kDa. Minimul de 55,37 kDa și 58,1 kDa posedă produsul de expresie a genei *Rf1b* de la *Oryza sativa* și *ALDH* de la *Arabidopsis thaliana*. Abaterea maximală pozitivă a fost înregistrată la produsul de expresie a genei *Rf1c* și *Rf1* de la *Oryza sativa* și la proteina restauratoare de la *Raphanus sativus* cu niște indici de 87,7 kDa, 87,6 kDa și 76,5 kDa.

Analiza *punctului izoelectric* al produșilor de expresie a genelor *ASC* a semnalat o abatere cuprinsă în limitele 4,82-9,87. Investigarea computerizată a acestei variabile, la proteinele restauratoare, a scos în evidență valori intermediare de 6,87, ceea ce reprezintă cu 1,49 unități mai puțin față de valorile medii ale proteinelor *ASC*. Extremele minime au fost înregistrate la produșii de expresie a genelor *orf687* de la *Raphanus sativus* – 5,23 și *Rf1a* de la *Oryza sativa* – 6,25. Maximul valorilor de 8,36 și 8,22 a fost observat la produșii de expresie a genelor *RfPPR591* de la *Petunia hzbrida* și *Rf1b* de la *Oryza sativa*.

Studiul detaliat al *hidrofobității proteinelor* a arătat că la ambele grupe *ASC* și *Rf* predomină catenele hidrofobe cu o valoare medie de 0,5, față de cele hidrofile cu o valoare medie de 0,23. La proteinele specifice genotipurilor androsterile valorile maxime ale catenelor hidrofobe de 0,66 și 0,63 au fost înregistrate la *ORF67* de la *Zea mays* și, respectiv, la *ORF79* de la *Oryza sativa*. La proteinele restauratoare de fertilitate analiza catenelor hidrofobe arată o abatere maximă de 0,07 și 0,06 la aldehydohidrogenaza de la *Oryza sativa* și la *ALDH2B* de la *Zea mays*.

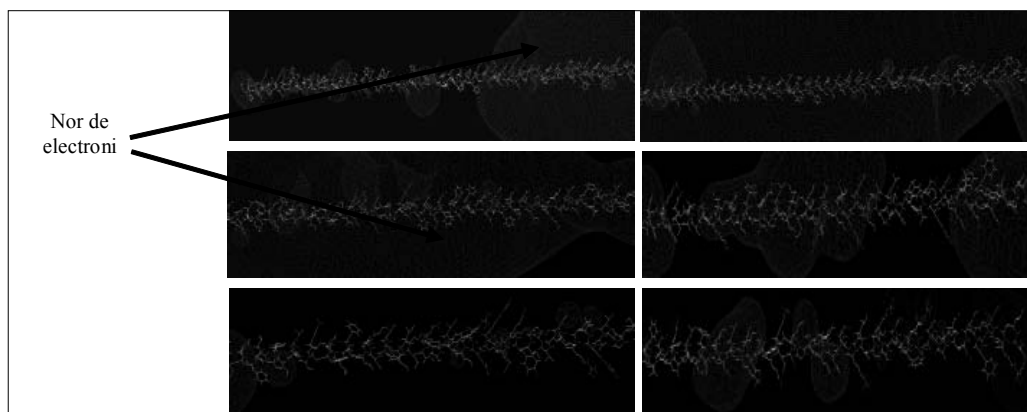
Analiza complexului proteic din cadrul grupului *ASC* a sugerat faptul că *potențialul electrostatic* e puțin deplasat în favoarea sarcinilor pozitive, cu o valoare medie de 0,11 față de 0,9 ale celor negative. Maximurile

sarcinilor negative predomină la proteina *ORF67* de la *Brassica oleracea* – 0,22, la proteina *ORF138* de la *Raphanus sativus* – 0,15 (Fig.1) și la produsul de expresie a genei *orf209* de la *Phaseolus vulgaris* – 0,14.



**Fig.1.** Potențialul electrostatic al proteinei *ORF138* de la *Raphanus sativus* modelat cu ajutorul programei Swiss Pdb Viewer ver. 4.0.1 [13].

Studiul proteinelor restauratoare de fertilitate indică faptul că atât sarcinile pozitive, cât și cele negative au o distribuție înaltă, cu un indice de 0,1. Numărul maxim de sarcini negative –0,13 a fost observat la produsul de expresie a genei *orf687* de la *Raphanus sativus* (Fig.2) și la produșii de expresie a genelor *Rf4* de la *Oryza sativa* și *RfPPR592* de la *Petunia hybrida*. Iar maxima sarcinilor pozitive este prezentă la produșii de expresie a genelor *RfPPR591* și *RfPPR592* de la *Petunia hybrida*.



**Fig.2.** Potențialul electrostatic la produsul de expresie a genei *orf687* de la *Raphanus sativus* modelat cu ajutorul programei Swiss Pdb Viewer ver. 4.0.1 [13]: roșu – nor de electroni; albastru – deficit de electroni. Notă: desenele reprezintă o continuitate de la stânga la dreapta.

Analiza **frecvenței aminoacizilor** la eșantionul compus din 24 de proteine androsterile a arătat că în majoritatea cazurilor predomină leucina, opozabil triptofanului cu cel mai mic procentaj al distribuției cantitative. Studiul proteinelor restauratoare de fertilitate a permis diferențierea acestora în două grupuri mari. În primul sunt incluse aldehiddehidrogenazele, unde dintre toți aminoacizii procentajul maximal îl posedă alanina, care, la rândul său, induce un indice hidrofobic mai mare întregii proteine, iar în al doilea grup predomină leucina.

Investigarea **structurii secundare** a proteinelor, care este în strânsă interdependență cu funcția acestora, a arătat în majoritatea cazurilor o predominanță de  $\beta$ -structuri la proteinele specifice genotipurilor androsterile. Prevalarea  $\alpha$ -helixurilor a fost întâlnită doar la produsul de expresie a genei *orfH522* de la *Helianthus annuus*, la proteinele *ORF209* de la *Phaseolus vulgaris* și *ORF108* de la *Brassica juncea*. Un număr egal de  $\alpha$ -helixuri și  $\beta$ -structuri a fost constatat pentru proteinele *ORF98*, *ORF239* de la *Phaseolus vulgaris* și *ORF222* de la *Brassica rapa* și *Brassica napus*.

Analiza produșilor de expresie a genelor *Rf* a permis divizarea acestora în trei grupuri mari. În primul sunt incluse proteinele genelor *Rf1*, *Rf1a*, *Rf1b* și *Rf1c* de la *Oryza sativa*, produsul de expresie a genei *rf2b* de la *Zea mays ASC-T*, proteinele *RfPPR591* (Fig.3) și *RfPPR592* de la *Petunia hybrida* și aldehiddehidrogenaza de la *Arabidopsis thaliana*, unde o predominanță majoră au  $\alpha$ -helixurile (Fig.3).



**Fig.3.** ALDH de la *Arabidopsis thaliana* și produsul de expresie a genei *RfPPR591* de la *Petunia hybrida* modelate cu ajutorul saitului [www.swissmodel.expasy.org](http://www.swissmodel.expasy.org) [17].

În al doilea grup intră aldehiddehidrogenazele de la *Oryza sativa* (Fig.4), *Nicotiana tabacum* și produsul de expresie a genei *rf2* de la *Zea mays* cu citoplasmă de tip T în care au fost remarcate proporții egale de  $\alpha$ -helixuri și  $\beta$ -structuri (Fig.4).



Fig.4. ALDH de la *Oryza sativa* și RF2 de la *Zea mays*, restaurator pentru ASC-T modelate cu ajutorul saitului [www.swissmodel.expasy.org](http://www.swissmodel.expasy.org) [17].

Iar în al treilea grup este inclus doar produsul de expresie a genei *Rfo* de la *Brassica napus*, unde o predominanță majoră au  $\beta$ -structurile.

#### Stabilirea tipurilor de domenii la proteinele cercetate

Un aspect important, legat de cercetarea proteinelor, îl reprezintă studiul domeniilor proteice. Aplicarea programei Phobius [11] a scos în evidență cele mai elocvente deosebiri dintre grupurile studiate. Astfel, majoritatea proteinelor responsabile pentru manifestarea fenotipică a ASC-ului au domenii noncitoplasmatică, transmembranare și citoplasmatică. Excepție de la regulă fac proteinele ASC: ORF108 de la *Brassica juncea*, ORF67 de la *Brassica oleracea*, ORF98 și ORF209 de la *Phaseolus sp.* și ORF25 de la *Triticum aestivum* – cu domenii strict citoplasmatică ca și toate proteinele restauratoare de fertilitate.

#### Discuții

În urma analizei a 24 de proteine androsterile și a 13 proteine restauratoare de fertilitate, au fost observate o serie de deosebiri și unele asemănări între cele două grupuri studiate (Fig.5).

PROTEINE ASC	PROTEINE Rf
Predomină catene Hidrofobe	
Un echilibru între sarcinile pozitive și cele negative	
Masă mică	Masă mare
Punctul izoelectric la un pH mai alcalin	Punctul izoelectric la un pH mai acid
Dintre aminoacizi predomină Leucina	Dintre aminoacizi predomină Leucina, la ALDH predomină Alanina
Predomină $\beta$ -structurile	Predomină $\alpha$ -helixurile

Fig.5. Deosebiri și asemănări între proteinele ambelor grupe studiate.

Masele moleculare mari ale proteinelor restauratoare de fertilitate și distribuția înaltă atât a sarcinilor pozitive, cât și a celor negative ce influențează proprietățile fizicochimice ale proteinelor, precum și interacțiunile dintre ele [5,6], ar putea fi un indiciu de complexitate structurală și funcțională a acestor biomolecule.

Totodată, proteinele androsterile, deși posedă dimensiuni mai mici, au înregistrat o predominanță de  $\beta$ -structuri opozabil proteinelor restauratoare de fertilitate cu dimensiuni mai mari, unde prevalează  $\alpha$ -helixurile. Acest fenomen poate fi explicat prin faptul că o structură biologică complicată, dacă ar fi formată din subunități foarte complexe, ar avea un număr foarte mare de probabilități de interacțiune dintre subunități, ceea ce la un moment dat ar duce la confruntarea între ele.

Prevalarea domeniilor transmembranare la majoritatea proteinelor androsterile poate fi un argument în favoarea concepțiilor din literatură precum că acestea sunt polipeptide membranare (cu referire la porumbul cu citoplasmă de tip T [3]; sfecla de zahăr [4]; petunie [8]; floarea - soarelui [2] etc. ).

Studiul combinat dintre programa Phobius [11] și Swiss Pdb Viewer ver. 4.0.1. [13] permite să observăm că capătul inițial de 23 de aminoacizi ai produsului de expresie a genei *orfH522* de la *Helianthus annuus* are un

domeniu noncitoplasmatic cu un potențial electrostatic echilibrat dintre sarcinile pozitive și cele negative. În partea centrală a acestuia (între 24 și 41 de aminoacizi) există, însă, un domeniu transmembranar în care lipsesc sarcinile electrice. Iar partea posterioară a proteinei (42-174 de aminoacizi cu un domeniu noncitoplasmatic) are un conglomerat de sarcini pozitive. Distribuția potențialului electrostatic pe fața internă și externă a membranei mitocondriale ne ajută să intuim repartizarea spațială la produsul de expresie a genei *orfH522* de la floarea-soarelui (Fig.6). Și deoarece studiul liniilor materne arată că proteina *ASC* cauzează expulzarea citocromului *c* din mitocondrie în citosol [1], putem intui că această proteină ar putea funcționa similar unei pompe ionice.

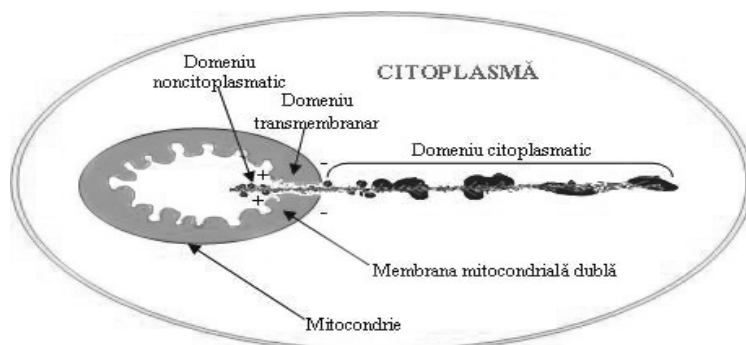


Fig.6. Modelul ipotetic privind distribuția spațială a produsului de expresie a genei *orfH522* de la *Helianthus annuus*.

### Concluzii

1. În urma cercetărilor am relevat că grupurile de proteine *ASC* și *Rf* au înregistrat o serie de deosebiri majore în ceea ce privește masa moleculară, punctul izoelectric și structura secundară.
2. Predominanța domeniilor transmembranare la majoritatea proteinelor *ASC* și lipsa acestora la producția de expresie a genelor *Rf* denotă că la primul grup sunt atribuite proteinele membranare.
3. Repartizarea spațială a produsului de expresie a genei *orfH522* de la *Helianthus annuus* ne permite să postulăm că, acesta ar putea funcționa similar unei pompe ionice, care expulzează citocromul *c* din mitocondrie în citosol.

### Referințe:

1. Balk J., Leaver C.J. The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome c release // *Plant Cell*. – 2001. – Vol.13. – P.1803-1818.
2. Horn R., Hustedt J.E.G., Horstmeyer A., Hahnen J., Zetsche K., Friedt W. The CMS-associated 16 kDa protein encoded by *orfH522* in the PET1 cytoplasm is also present in other male-sterile cytoplasms of sunflower // *Plant Molecular Biology* – 1996. – Vol.30. – P.523-538.
3. Kaspi C.I., Siedow J.N. Cross-linking of the cms-T Maize Mitochondrial Pore-forming Protein URF13 by N,N'-Dicyclohexylcarbodiimide and Its Effect on URF13 Sensitivity to Fungal Toxins // *The Journal of Biological Chemistry*. – 1993. – Vol.1. – P.5828-5823.
4. Kitazaki K., Nomoto Y., Kubo T., Mikami T. Expression in plants of a CMS-associated gene, preSatp6, from the Owen cytoplasm of sugar beet // *Proceedings of the Japanese Society of Sugar Beet Technologists*. – 2005. – No 46. – P.31-36.
5. Nishikawa M., Hasegawa S., Yamashita F. et al. Electrical charge on protein regulates its absorption from the rat small intestine // *Am J. Physiol Gastrointest Liver Physiol*. – 2002. – Vol.282. – P.711-719.
6. Sheinerman F.B., Norel R., Honig B. Electrostatic aspects of protein-protein interactions // *Curr. Opin. Struct. Biol.* – 2000. – Vol.10. – P.153-159.
7. Watson J.D., Crick F.H.C. Molecular structure of nucleic acids // *Nature*. – 1953. – Vol.171. – P.737-738.
8. Wintz H., Chen H., Sutton C.A., Conley C.A., Cobb A., Ruth D., Hanson M.R. Expression of the CMS-associated *urfS* sequence in transgenic petunia and tobacco // *Plant Molecular Biology*. – 1995. – Vol.28. – P.83-92.

### Surse web

9. [www.clcbio.com](http://www.clcbio.com).
10. [www.ddbj.nig.ac.jp/searches-e.html](http://www.ddbj.nig.ac.jp/searches-e.html)
11. [www.ebi.ac.uk/phobius/](http://www.ebi.ac.uk/phobius/).
12. [www.expasy.org/](http://www.expasy.org/)
13. [www.expasy.org/spdbv/](http://www.expasy.org/spdbv/).
14. [www.life.illinois.edu/z-huang/Genebank98.html](http://www.life.illinois.edu/z-huang/Genebank98.html)
15. [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/).
16. [www.publicsafety.gc.ca/prg/cor/tls/dna-eng.aspx](http://www.publicsafety.gc.ca/prg/cor/tls/dna-eng.aspx)
17. [www.swissmodel.expasy.org/woekspace/index.php](http://www.swissmodel.expasy.org/woekspace/index.php).