## **UNIVERSITATEA DE STAT DIN MOLDOVA**

Cu titlu de manuscris C.Z.U: 577.1./.2(043)

# CHERDIVARĂ ALA

# GLOBULINELE DE REZERVĂ 7S ȘI 11S DIN SEMINȚE: EVOLUȚIE, PROTEOLIZĂ, ALERGENICITATE

163.02 – Biochimie

Teză de doctor în științe biologice

Conducător științific:

ŞUTOV Andrei, doctor habilitat în științe biologice, profesor universitar

Autor:

Suto4 Adda

CHERDIVARĂ Ala

# CHIŞINĂU, 2020

© Cherdivară Ala, 2020

# CUPRINS

	ADNOTARE (română, engleză, rusă)	5
	LISTA TABELELOR	8
	LISTA FIGURILOR	9
	LISTA ABREVIERILOR	14
	INTRODUCERE	15
1.	GLOBULINELE DE REZERVĂ DIN SEMINȚE	22
	1.1. Structura globulinelor de rezervă din semințe	22
	1.2. Originea și evoluția globulinelor de rezervă	28
	1.3. Proteoliza globulinelor de rezervă și reglarea acesteia	36
	1.4. Globulinele de rezervă din semințe ca alergeni	42
	1.5. Concluzii la capitolul 1	47
2.	ORIGINEA ȘI EVOLUȚIA STRUCTURILOR GLOBULINELOR DE	
	REZERVĂ DIN SEMINȚE	48
	2.1. Metodele și strategia de cercetare	48
	2.2. Analiza originii și evoluției globulinelor de rezervă	50
	2.3. Succesiunea formării structurilor globulinelor 7S și 11S	62
	2.4. Controlul degradării globulinelor de rezervă în semințele germinative	64
	2.5. Concluzii la capitolul 2	68
3.	LEGITĂȚILE PROTEOLIZEI GLOBULINEI 11S DIN SEMINȚELE DE	
	GINKGO	69
	3.1. Materialul și metodele de cercetare	69
	3.2. Izolarea și purificarea ginnacinei	71
	3.3. Proteoliza și cinetica proteolizei ginnacinei	73
	3.4. Modelarea omoloagă a structurilor terțiare și cuaternare ale ginnacinei	76
	3.5. Concluzii la capitolul 3	88
	·	

4.	PROTEOLIZA CA MODALITATE DE REDUCERE A ALERGENICITĂȚII	
	GLOBULINELOR DE REZERVĂ DIN ARAHIDE	89
	4.1. Materialul și metodele de cercetare	89
	4.2. Izolarea și purificarea globulinei 7S din arahide	91
	4.3.Proteoliza limitată a Ara h1, globulina 78 din arahide	92
	4.4.Izolarea și purificarea globulinei 11S din arahide	99
	4.5.Proteoliza limitată a Ara h3, globulina 11S din arahide	101
	4.6.Imunoreactivitatea încrucișată a globulinelor de rezervă	112
	4.7.Concluzii la capitolul 4	116

CONCLUZII GENERALE ȘI RECOMANDĂRI	118
BIBLIOGRAFIE	120
ANEXE	130
Anexa 1. Act de implementare a rezultatelor științifice în instruire	130
DECLARAȚIA PRIVIND ASUMAREA RĂSPUNDERII	131
CV-ul AUTORULUI	132

#### ADNOTARE

#### la teza "**Globulinele de rezervă 7S și 11S din semințe: evoluție, proteoliză, alergenicitate**" prezentată de Cherdivară Ala în vederea solicitării gradului stiintific de doctor în stiinte biologice. Chisinău, 2020

Teza include introducere, patru capitole, concluzii generale, recomandări și lista lucrărilor citate. Lucrarea conține 116 pagini de text, 64 de figuri și 14 tabele. Referințele includ 120 de surse. Rezultatele obținute sunt publicate în 18 lucrări științifice (9 articole și 9 teze la conferințe).

**Cuvinte-cheie:** globuline de rezervă din semințe, legumine, viciline, germine, evoluție, ginkgo, oxalat-decarboxilaza, structurile primare și superioare ale proteinelor, mecanisme de proteoliză, proteoliză limitată, cinetica proteolizei, alergenicitate, arahide.

**Scopul lucrării:** Studierea legităților formării structurii conservative a globulinelor de rezervă din semințe în timpul evoluției, stabilirea mecanismului degradării proteolitice profunde a acestora și analiza perspectivelor de utilizare a proteolizei limitate a globulinelor de rezervă pentru a reduce nivelul de alergenicitate a acestora.

**Obiectivele lucrării:** 1) Analiza relațiilor filogenetice dintre globulinele de rezervă din semințe și proteinele structural omoloage ale procariotelor și eucariotelor. 2) Stabilirea etapelor succesive în formarea structurii globulinelor de rezervă, care determină funcția lor în semințe. 3) Studierea cineticii proteolizei globulinei de rezervă 11S din semințele de ginkgo *Ginkgo biloba*, unul dintre cei mai vechi reprezentanți ai spermatofitelor și descrierea mecanismului de reglare a degradării lor proteolitice profunde. 4) Studierea reacțiilor succesive a proteolizei limitate ale globulinelor 11S și 7S din arahide *Arachis hypogaea*, semințele cărora au un nivel ridicat de alergenicitate. 5) Stabilirea modificărilor structurale în globulinele 11S și 7S în timpul proteolizei și analiza posibilității de reducere a alergenicității globulinelor de rezervă din arahide prin proteoliza limitată.

Noutatea și originalitatea științifică a lucrării: Pentru prima dată, s-a stabilit că precursorii străvechi evolutivi ai globulinelor de rezervă din semințe sunt oxalat-decarboxilazele bacteriene. S-a arătat că precursorii străvechi vegetali ai globulinelor de rezervă din semințe sunt proteinele din algele verzi *Klebsormidium nitens*. S-a constatat că structurile primare și terțiare ale globulinelor 11S reflectă cele mai vechi trăsături ale proteinelor de rezervă din semințe. S-a demonstrat că proteoliza limitată a globulinei 11S din semințele de ginkgo inițiază degradarea sa masivă prin mecanismul "una câte una". Pe baza analizei structurilor cristaline ale globulinelor 11S, este prezentat rolul reglator al proteolizei limitate în degradarea masivă a acestor proteine. S-au studiat legitățile proteolizei limitate ale globulinelor 11S și 7S din semințele de arahide. S-a stabilit că proteoliza limitată duce la îndepărtarea regiunilor structurii primare ale globulinelor 11S și 7S din arahide, care determină alergenicitatea lor.

**Rezultatul obținut**, care contribuie la soluționarea unei probleme științifice importante constă în stabilirea originii globulinelor de rezervă din semințe din oxalat-decarboxilazele bacteriene și rolului de reglare al proteolizei limitate în degradarea lor profundă. S-a demonstrat că proteoliza limitată a proteinelor de rezervă din semințe duce la eliminarea determinanților antigenici (epitopii IgE) din structurile lor primare, ceea ce este promițător din punct de vedere al creșterii siguranței produselor alimentare, care conțin proteine de rezervă din semințe.

Semnificația teoretică a lucrării: S-a stabilit originea globulinelor de rezervă din semințe din oxalat-decarboxilazele bacteriene și nu din sferulinele Amoebozoa, după cum s-a presupus anterior. S-a stabilit rolul proteolizei limitate în reglarea degradării proteolitice masive a globulinei 11S din ginkgo, unul dintre cei mai vechi reprezentanți ai spermatofitelor. Rezultatele obținute sugerează un mecanism comun pentru reglarea proteolizei proteinelor de rezervă, caracteristic pentru întreaga familie de globuline 11S din semințe.

**Semnificația aplicativă a lucrării:** Sunt prezentate perspectivele proteolizei limitate ca o modalitate de a reduce semnificativ nivelul de alergenicitate al proteinelor de rezervă din semințele de arahide, precum și a mai multor altor plante.

**Implementarea rezultatelor științifice:** Rezultatele expuse în lucrare reprezintă material științifico-didactic pentru cursurile de biochimie și proteomică.

#### SUMMARY

#### of the thesis "**7S and 11S seed storage globulins: evolution, proteolysis, allergenicity**" presented by Cherdivară A.M. for the scientific degree of Doctor in Biology, Chisinau, 2020

The thesis consists of introduction, four chapters, general conclusions and recommendations, and a list of references. The work contains 116 pages of the text, 64 figures, 14 tables and the list of references that includes 120 sources. The obtained results were published in 18 scientific papers (9 articles and 9 reports at conferences).

**Key words:** seed storage globulins, legumin, vicilin, germin, oxalate decarboxylase, evolution, protein primary and higher order structures, proteolysis mechanisms, limited proteolysis, proteolysis kinetics, allergenicity, ginkgo, peanut.

Aim of the study: Investigation of the regularities of formation during evolution of the conserved structure of seed storage globulins, establishment of the mechanism of their deep proteolytic degradation and elucidation of the prospects for the use of limited proteolysis of storage globulins to reduce the level of their allergenicity.

**The objectives of the study.** 1) To analyze phylogenetic relationships between sequences of seed storage globulins and related proteins of prokaryotes and eukaryotes. 2) To establish successive stages in the evolution of the structure of storage globulins that determine their function. 3) To study the kinetics of proteolysis of the storage 11S globulin of ginkgo *Ginkgo biloba*, one of the most ancient representatives of spermatophytes, and to describe the mechanism of regulation of deep proteolytic degradation of 11S globulins. 4) To study the successive reactions of limited proteolysis of 11S and 7S globulins from peanut *Arachis hypogaea*, whose seeds are known for a high level of allergenicity. 5) To establish structural changes in 11S and 7S globulins during proteolysis and assess the possibility of reducing the allergenicity of peanut storage globulins by limited proteolysis.

Scientific novelty and originality of the work. It was established for the first time that bacterial oxalate decarboxylases represent statistically reliable most ancient precursors of seed storage globulins. It was demonstrated that proteins from the green algae *Klebsormidium nitens* represent the most ancient plant precursors of seed storage globulins. It was shown that the primary and higher order structures of 11S globulins reflect the most ancient features of seed storage globulins. It was demonstrated that the limited proteolysis of 11S globulin from ginkgo seeds initiates massive degradation of the protein via one-by-one mechanism. Based on the analysis of crystal structures of 11S globulins, the regulatory role of the limited proteolysis in the massive degradation of these proteins was shown. Regularities of limited proteolysis leads to the removal of the primary structure regions of peanut 11S and 7S globulins, which determine their allergenicity.

The obtained result, which contributes to the solution of an important scientific problem, is to establish the origin of seed storage globulins from bacterial oxalate decarboxylases and the regulatory role of limited proteolysis in their deep degradation. It has been shown that limited proteolysis of seed storage globulins leads to the removal of antigenic determinants (IgE epitopes) from their primary structures, which is promising from the point of view of increasing the safety of food products containing seed proteins.

**Theoretical significance of the work.** The origin of seed storage globulins from bacterial oxalate decarboxylases, rather than from Amoebozoa spherulins as assumed previously, was shown statistically faithfully. The role of limited proteolysis in the regulation of massive proteolytic degradation of 11S globulin from ginkgo, one of the oldest representatives of seed plants, has been established. The results obtained suggest a common mechanism for the regulation of proteolysis of storage proteins characteristic of the entire family of 11S seed globulins.

**Practical significance of the work.** The availability of the limited proteolysis as a means of a significant decrease of the allergenicity level of storage globulins from pumpkin seeds, as well as from seeds of other plants, were shown.

**Implementation of scientific results.** The results presented in the paper represent scientific-didactic material for the courses of biochemistry and proteomics.

#### АННОТАЦИЯ

#### диссертации Кердиварэ А.М. «Запасные 7S и 11S глобулины семян: эволюция, протеолиз, аллергенность», представленной на соискание ученой степени доктора биологических наук, Кишинев, 2020

Диссертация состоит из введения, четырех глав, общих выводов, рекомендаций и списка цитируемой литературы. Работа состоит из 116 страниц текста, 64 рисунков и 14 таблиц. Список литературы включает 120 источников. Полученные результаты опубликованы в 18 научных работах (9 статей и 9 докладов на конференциях).

**Ключевые слова:** запасные глобулины семян, легумин, вицилин, джермин, оксалатдекарбоксилаза, эволюция, первичная и высшие структуры белков, механизмы протеолиза, ограниченный протеолиз, кинетика протеолиза, аллергенность, гинкго, арахис.

**Цель исследования:** Исследование закономерностей формировании в ходе эволюции консервативной структуры запасных глобулинов семян, установление механизма их глубокой протеолитической деградации и выяснение перспектив использование ограниченного протеолиза запасных глобулинов для снижения уровня их аллергенности.

Задачи исследования: 1) Провести анализ филогенетических взаимоотношений между последовательностями запасных глобулинов семян и родственных белков прокариот и эукариот. 2) Установить последовательные этапы эволюции структуры запасных глобулинов, определяющие их функцию. 3) Исследовать кинетику протеолиза запасного 11S глобулина гинкго *Ginkgo biloba*, одного из древнейших представителей сперматофитов и описать механизм регуляции глубокой протеолитической деградации 11S глобулинов. 4) Изучить последовательные реакции ограниченного протеолиза 11S и 7S глобулинов арахиса *Arachis hypogaea*, семена которого известны высоким уровнем аллергенности. 5) Установить структурные изменения 11S и 7S глобулинов в ходе протеолиза и оценить возможности снижения аллергенности запасных глобулинов арахиса методом ограниченного протеолиза.

Научная новизна и оригинальность работы. Впервые установлено, что наиболее древними эволюционными предшественниками запасных глобулинов семян являются бактериальные оксалатдекарбоксилазы. Показано, что наиболее древними растительными предшественниками запасных глобулинов семян являются белки зеленой водоросли *Klebsormidium nitens*. Установлено, что первичная структура 11S глобулинов отражает наиболее древние признаки запасных белков семян. Показано, что ограниченный протеолиз 11S глобулина семян гинкго инициирует его массированную деградацию по поодиночному механизму. На основе анализа кристаллических структур 11S глобулинов показана регуляторная роль ограниченного протеолиза в массированной деградации этих белков. Изучены закономерности ограниченного протеолиза 11S и 7S глобулинов семян арахиса. Установлено, что ограниченный протеолиз приводит к удалению областей первичной структуры 11S и 7S глобулинов арахиса, определяющих их аллергенность.

**Полученный результат,** который способствует решению важной научной проблемы, заключается в установлении происхождения запасных глобулинов семян из бактериальных оксалат-декарбоксилаз и регуляторной роли ограниченного протеолиза в их глубокой деградации. Показано, что ограниченный протеолиз запасных глобулинов семян приводит к удалению из их первичных структур антигенных детерминант (IgE-эпитопов), что является перспективным с точки зрения повышения безопасности пищевых продуктов, содержащих белки семян.

**Теоретическая значимость работы:** Установлено происхождение запасных глобулинов семян из бактериальных оксалат-декарбоксилаз, а не из сферулинов Amoebozoa, как ранее предполагалось. Установлена роль ограниченного протеолиза в регуляции массированной протеолитической деградации 11S глобулина гинкго, одного из древнейших представителей семенных растений. Полученные результаты позволяют предполагать общность механизма регуляции протеолиза запасных белков, характерного для всего семейства 11S глобулинов семян.

**Практическая значимость работы:** Показана перспективность метода ограниченного протеолиза как способа существенного снижения уровня аллергенности запасных белков семян арахиса, а также ряда других растений.

**Внедрение научных результатов**. Результаты, изложенные в работе, представляют собой научно-образовательный материал для курсов биохимии и протеомики.

### LISTA TABELELOR

Tabelul 1.1.	Globulinele 7S cu structura terțiară cunoscută	23	
Tabelul 1.2.	Globulinele 11S cu structura terțiară cunoscută		
Tabelul 1.3.	Caracteristicile principale ale reacțiilor idealizate ale proteolizei		
	limitate și "una câte una" a proteinelor native	35	
Tabelul 1.4.	Poziția legăturilor peptidice, scindate în timpul proteolizei limitate, în		
	structura terțiară a subunităților globulinei 7S din soia (GLYma) și din		
	fasole (PHAvu).	38	
Tabelul 1.5.	Poziția legăturilor peptidice, scindate la proteoliză limitată, în		
	structura terțiară a subunităților globulinelor 11S din soia (GLYma		
	G1-G5), din dovleac Cucurbita maxima (CUCma), din floarea-		
	soarelui Helianthus annuus (HELan), din ovăz Avena sativa (AVEsa),		
	din cedru Pinus sibirica (PINsi)	39	
Tabelul 1.6.	Subunitățile globulinelor 7S care sunt alergeni	41	
Tabelul 1.7.	Subunitățile globulinelor 11S, care sunt alergeni	43	
Tabelul 1.8.	Subunitățile de tipul I-IV, ce diferă după masa moleculară (kDa), din		
	molecula heterohexamerică a globulinei 11S din arahide	44	
Tabelul 2.1.	Taxonomia organismelor, ale căror proteine sunt folosite pentru a		
	analiza relațiile evolutive dintre globulinele de rezervă și germine	50	
T <b>abelul 3.1.</b>	Analiza oportunității utilizării ca model structurile cristaline ale		
	globulinelor 11S pentru modelarea structurii terțiare și cuaternare a		
	ginnacinei	75	
Tabelul 3.2.	Accesibilitatea la solvent și legăturile de hidrogen intersubunitare		
	în oligomerii globulinelor 11S	82	
Tabelul 4.1	Legăturile peptidice scindate de papaină din subunitățile Ara h1	92	
Tabelul 4.2.	Masele moleculare (kDa) a celor patru tipuri de subunități ale		
	globulinei 11S din arahide	103	
Tabelul 4.3.	Evaluarea cantitativă a imunoreactivității încrucișate a globulinelor de		
	rezervă în regiunile omoloage ale structurilor lor.	112	

## LISTA FIGURILOR

Figura 1.1.	Diagrama panglică a structurii domeniului N-terminal al subunității $\alpha'$		
	al globulinei 7S din soia pdb 1uik	22	
Figura 1.2.	Secvența de aminoacizi a subunității $\alpha$ ' a globulinei 7S din soia		
	pdb 1uik	22	
Figura 1.3.	Structura cuaternară a globulinei 7S din soia 1ipk	23	
Figura 1.4.	Diagrama panglică a structurii terțiare a domeniului N-terminal al		
	subunității A3B4 al globulinei 11S din soia pdb 10d5	24	
Figura 1.5.	Secvența de aminoacizi a subunității A3B4 a globulinei 11S din soia,		
	glicinina pdb 1od5	25	
Figura 1.6.	Structura cuaternară a globulinei 11S din soia 1fxz	26	
Figura 1.7.	Diagramele panglică ale structurilor domeniilor C-terminale a		
	globulinei 11S din soia 1fxz (a) și a globulinei 7S Canavalia		
	ensiformis 2cau (b) suprapuse spațial, dar arătate separat	28	
Figura 1.8.	Alinierea structurală a secvențelor de aminoacizi ale globulinei 11S		
	din soia 1 fxz și globulinei 7S Canavalia ensiformis 2cau	29	
Figura 1.9.	Diagramele panglică ale structurii terțiare al domeniului N-terminal a		
	globulinei 7S Canavalia ensiformis 2cau și germinei pdb 1fi2 din orz		
	suprapuse spațial, dar prezentate separat	30	
Figura 1.10.	Alinierea secvențelor de aminoacizi ale domeniului N-terminal al		
	globulinei 7S Canavalia ensiformis 2cau și al zonei modulului		
	structural β-barrel-α-helixuri al germinei 1fi2	31	
Figura 1.11.	Diagramele panglică ale structurilor terțiare ale domeniilor C-		
	terminale ale globulinei 11S din soia pdb lod5 (A) și OD		
	cianobacteriilor Synechocystis sp. pdb 2vqa (B) suprapuse spațial, dar		
	prezentate separat	32	
Figura 1.12	Alinierea secvențelor de aminoacizi ale subunităților globulinei 11S		
	din soia 10d5 și OD cianobacteriei Synechocystis sp. pdb 2vqa	33	
Figura 1.13	Analiza filogenetică a secvențelor de aminoacizi ale globulinelor de		
	rezervă (și a precursorilor acestora din plantele cu spori), germinelor		
	(și precursorii lor sferulinelor) și OD	34	
Figura 1.14	Poziția epitopilor IgE în structura terțiară a subunităților globulinei 7S		
	din arahide, Ara h1 [78]	42	

Figura 1.15	Poziția epitopilor IgE în structurile terțiare ale subunităților		
	globulinelor 11S din arahide, Ara h3 [82] și soia Gly m G1 [84] și		
	Gly m G2 [85]	44	
Figura 2.1.	Arborele fără rădăcină, care descrie relația filogenetică dintre		
	clusterele globulinelor de rezervă, OD și germinelor	49	
Figura 2.2.	Alinierea secvențelor de aminoacizi ale sferulinei aaa29979 din		
	Amoebozoa Physarum polycephalum (PHYpo) și ale proteinei OD-		
	similare cam64834 din actinobacteria Mycobacterium abcesus		
	(MYCab)	51	
Figura 2.3.	Căile independente ale evoluției globulinelor de rezervă din semințe		
	și a germinelor	52	
Figura 2.4.	Alinierea secvențelor de aminoacizi ale proteinei gaq91246 din		
	Klebsormidium nitens și globulinei 11S pdb 3qac din Amaranthus		
	hypochondriacus	54	
Figura 2.5.	Diagramele panglică ale structurilor terțiare ale domeniilor C-		
	terminale ale proteinei gaq91246 din alga verde Klebsormidium		
	nitens (a, model) și globulinei 11S pdb 3qac din Amaranthus		
	hypochondriacus (b), suprapuse spațial, dar prezentate separat	55	
Figura 2.6.	Structura cuaternară model a proteinei leguminosimilară gaq91246		
	din alga verde Klebsormidium nitens	56	
Figura 2.7.	Alinierea secvențelor de aminoacizi, utilizate pentru construirea		
	arborelui evolutiv al superfamiliei de proteine de rezervă (partea inițială).	57	
Figura 2.7.	Alinierea secvențelor de aminoacizi, utilizate pentru construirea arborelui		
	evolutiv al superfamiliei de proteine de rezervă (pozițiile finale)	58	
Figura 2.8.	Calea evolutivă a globulinelor de rezervă din semințe	59	
Figura 2.9.	Compararea structurilor terțiare (coloane albe) și primare (coloane		
	negre) ale subunităților globulinelor de rezervă și a OD 2vqa din		
	cianobacterii Synechocystis sp	61	
Figura 2.10.	Diagramele panglică ale structurilor subunităților A și B în		
	componența trimerilor OD 2vqa din cianobacterii Synechocystis sp.		
	(a), leguminei 3fz3 Prunus dulcis (b) și vicilinei 1uik Glycine max		
	(c), suprapuse spațial, dar prezentate separat	62	

Fugura 2.11.	Glicinina 1 fxz din soia: regiunile structural organizate ale α-catenei	
	(a) și diagrama panglică a structurii ei terțiare (b)	65
Figura 3.1.	SDS-PAGE proteinelor din extractul brut din semințele de ginkgo	
	(pistele 1 și 5) și fracțiilor precipitate cu sulfat de amoniu	70
Figura 3.2.	Cromatografia pe coloana cu fenil-Sepharose CL-4B a ginnacinei,	
	parțial purificată prin salifiere	70
Figura 3.3.	SDS – PAGE preparatului de globulină 11S, parțial purificat prin	
	salifiere cu sulfat de amoniu (pistele 1 și 4) și fracțiilor eluate la	
	cromatografia fenil-Sepharose	71
Figura 3.4.	SDS-PAGE ginnacinei intacte și produselor ei de proteoliză	72
Figura 3.5.	Gel-filtrarea pe coloana cu Sephacryl S-300 a produselor de hidroliză	
	a ginnacinei după 24 ore de reacție	72
Figura 3.6.	Cinetica proteolizei ginnacinei	73
Figura 3.7.	Relațiile filogenetice dintre ginnacină și globulinele 11S cu structura	
	terțiară cunoscută	74
Figura 3.8.	Diagramele panglică ale structurilor $\alpha$ -catenelor ginnacinei (a) și	
	globulinei 11S pdb 3qac din Amaranthus hypochondriacus (b)	
	suprapuse spațial, dar prezentate separat	76
Figura 3.9.	Alinierea secvențelor de aminoacizi ale $\alpha$ -catenelor ginnacinei (G) și	
	globulinei 11S din amarant Amaranthus hypochondriacus (A)	77
Figura 3.10.	Alinierea secvențelor de aminoacizi ale $\alpha$ -catenelor globulinelor 11S	
	din dovleac 2evx, din soia 10d5 și ginnacinei (GINbi)	78
Figura 3.11.	Compararea rezultatelor modelării structurilor secundare ale a-	
	catenelor ginnacinei	79
Figura 3.12.	Compararea structurilor secundare, prezente real în structura cristalină	
	pdb 3qac a globulinei 11S din Amaranthus hypochondriacus (S, $\beta$ -	
	strendurile, H, $\alpha$ -helixurile) cu rezultatele modelării structurilor	
	secundare ale acestei proteine, utilizând programul SSPRO 8	80
Figura 3.13.	Legăturile de hidrogen (linii punctate) responsabile de interacțiunea	
	dintre subunitățile A și B în structura cuaternară model a ginnacinei.	83
Figura 3.14.	Numărul legăturilor de hidrogen dintre subunitățile A și B	84
Figura 3.15.	Legăturile hidrogen (linii punctate) între restul de aminoacid Gln155 al	
	subunității A și resturile Asn389, Gln426 și Thr446 ale subunității B în	
	structura cuaternară a globulinei 11S 2d5h din soia	84

Figura 3.16.	Restul de aminoacid Gln155 din regiunea dintre $\beta$ -strendurile I și J ale		
	subunității A unit prin legături de hidrogen cu resturile Asn389,		
	Gln426 și Thr446 din structura subunității B a globulinei 11S 2d5h		
	din soia	85	
Figura 3.17.	Alinierea secvențelor de aminoacizi ale globulinelor 11S cu structura		
	cuaternară cunoscută	86	
Figura 4.1.	SDS-PAGE extractului brut din semințele de arahide (pista 1) și		
	fracțiile precipitate cu sulfat de amoniu	89	
Figura 4.2.	Cromatografia pe coloana cu fenil-Sepharose CL-4B (a) și SDS-		
	PAGE (b) a preparatului de globulină 7S din arahide, parțial purificat		
	prin salifiere cu sulfat de amoniu	90	
Figura 4.3.	Diagramele panglică ale structurii terțiare ale domeniilor N-terminal (a)		
	și C-terminal (b) în subunitățile globulinei 7S, Ara h1, din arahide	91	
Figura 4.4.	SDS-PAGE a alergenului 7S, Ara h1, în timpul hidrolizei cu papaină.	92	
Figura 4.5.	Schema ipotetică de formare a fragmentelor Ara h1 în timpul		
	hidrolizei cu papaină	93	
Figura 4.6.	Secvența de aminoacizi a globulinei 7S, Ara h1, din arahide	95	
Figura 4.7.	SDS – PAGE a extractului brut din semințele de arahide și fracțiile		
	precipitate cu sulfat de amoniu	98	
Figura 4.8.	Cromatografia pe fenil-Sepharose CL-4B a preparatului de globulină		
	11S, parțial purificat prin salifiere cu sulfatul de amoniu (fracția 30-		
	40% de saturație)	99	
Figura 4.9.	SDS-PAGE a preparatului de globulină 11S, parțial purificat prin		
	salifiere cu sulfatul de amoniu și a fracțiilor eluate prin cromatografie		
	pe fenil-Sepharose cu NaCl 1M	99	
Figura 4.10.	Secvența de aminoacizi și structurile secundare ale $\alpha$ -catenei		
	globulinei 11S din arahide (pdb 3c3v)	100	
Figura 4.11.	Diagrama panglică a structurii terțiare a globulinei 11S din arahide		
	pdb 3c3v	101	
Figura 4.12.	Diagramele panglică a structurii cuaternare ale globulinei 11S din		
	arahide 3c3v	102	
Figura 4.13.	Resturile de aminoacizi cu accesibilitate sporită la solvent (ASA>90%)		
	în structura hexamerică model a globulinei 11S din arahide, Ara h3	102	

<b>4.14.</b> SDS-PAGE globulinei 11S din arahide în timpul proteolizei cu			
papaină în prezența ME	104		
Proteoliza globulinei 11S din arahide cu papaina	104		
SDS-PAGE, în prezența (+ME) și în absența (-ME) 2-ME, globulinei			
11S din arahide în timpul proteolizei cu tripsină	106		
Analiza cantitativă a proteolizei globulinei 11S din arahide cu tripsină	107		
Schema secvențelor $\alpha$ -catenelor în subunitățile 3c3v și ABL14270 ale			
globulinei 11S din arahide	108		
Imunoreactivitatea încrucișată a epitopilor IgE ai globulinelor 11S din			
arahide, Ara h3 (epitopul 2) și din soia Gly m G1 (epitopul 1),			
precum și ai globulinei 7S din arahide, Ara h1 (epitopul 11)	111		
Imunoreactivitatea încrucișată a globulinelor de rezervă din semințe			
în regiunile omoloage ale structurilor primare din regiunea $\alpha$ -	113		
helixurilor h1 și h2 ale domeniilor N-terminale			
	SDS-PAGE globulinei 11S din arahide în timpul proteolizei cu papaină în prezența ME Proteoliza globulinei 11S din arahide cu papaina SDS-PAGE, în prezența (+ME) și în absența (-ME) 2-ME, globulinei 11S din arahide în timpul proteolizei cu tripsină Analiza cantitativă a proteolizei globulinei 11S din arahide cu tripsină Schema secvențelor $\alpha$ -catenelor în subunitățile 3c3v și ABL14270 ale globulinei 11S din arahide Imunoreactivitatea încrucișată a epitopilor IgE ai globulinelor 11S din arahide, Ara h3 (epitopul 2) și din soia Gly m G1 (epitopul 1), precum și ai globulinei 7S din arahide, Ara h1 (epitopul 11) Imunoreactivitatea încrucișată a globulinelor de rezervă din semințe în regiunile omoloage ale structurilor primare din regiunea $\alpha$ - helixurilor h1 și h2 ale domeniilor N-terminale		

#### LISTA ABREVIERILOR

- ASA Suprafața accesibilă la solvent unui rest de aminoacid (Accessible Surface Area)
- ATA Acid tricloracetic
- BLAST Instrumentul de bază de căutare a alinierii locale (Basic Local Alignment Search Tool)
- EDTA Acid etilendiaminotetraacetic (Ethylenediaminetetraacetic acid)
  - E-64 Trans-epoxysuccinnyl-L-leucylamido-(4-guanidino)butan
  - FAO Organizația pentru Alimentație și Agricultură (Food and Agriculture Organization)
  - Ig E Imunoglobulina E
  - kDa Kilo Dalton
  - ME 2-mercaptoetanol
- NCBI Centrul Național pentru Informații Biotehnologice (National Center for Biotechnology Information)
  - OD Oxalat-decarboxilaze
- PAGE Electroforeza în gel de poliacrilamidă (Polyacrylamide Gel Electrophoresis)
- PD-indice Indice de similaritate peptidică bazat pe proprietatea a două secvențe (Property-Based Peptide Similarity Index for Two Sequences)
  - RMSD Deviația rădăcină-medie-pătrată a pozițiilor atomice (Root-Mean-Square Deviation)
    - SA Sulfat de amoniu
  - SDAP Baza de date structurale ale proteinelor alergenice (Structural Database of Allergenic Proteins)
    - SDS Dodecilsilfat de sodiu (Sodium Dodecyl Sulfate)
  - TEMED Tetrametildiamina
    - Tris Hidroximetilaminometan

#### INTRODUCERE

Actualitatea și importanța temei abordate. Semințele sunt o sursă importantă de proteine alimentare. Conform datelor Organizației pentru Alimentație și Agricultură a ONU (FAO - Food and Agriculture Organization) rația alimentară a oamenilor la nivel mondial conține 70% proteine vegetale și doar 30% – proteine animale. Partea principală a proteinelor vegetale alimentare sunt proteinele de rezervă din semințe. Numărul mare de proteine de rezervă din semințe este inclus în două familii – globuline 7S (viciline) și globuline 11S (legumine) [1].

Globulinele 7S și 11S din semințe reprezintă o rezervă de proteine macromoleculare, care se formează în timpul maturizării semințelor, iar la sfârșitul perioadei de repaus, indefinit de lungă, sunt mobilizate pentru a hrăni germenul. În ciuda primitivități aparente a funcției, globulinele de rezervă din semințele tuturor spermatofitelor sunt foarte conservative, ceea ce indică condiționalitatea funcțională a structurii lor. Spre deosebire de alte proteine din semințe, al căror metabolism constă în sinteza și degradarea, care decurg paralel, în globulinele de rezervă 7S și 11S, aceste procese sunt separate în timp. Prin urmare, structura globulinelor de rezervă formată în timpul evoluției este protejată în semințele germinative de degradarea prematură și dobândește capacitatea de a regla mobilizarea proteolitică în timpul germinării semințelor și dezvoltării germenului [2].

Componența aminoacidică și proprietățile funcționale (tehnologice) ale globulinelor de rezervă native din semințe sunt departe de a fi întotdeauna satisfăcătoare, ceea ce limitează utilizarea directă a acestora în produsele alimentare și ca componente ale formelor artificiale de alimente bogate în proteine. Mai mult decât atât, conform SDAP (baza de date structurale a proteinelor alergene) [3], globulinele de rezervă din semințele multor plante sunt alergeni. Încercările de a utiliza tehnicile inginerie genice pentru îmbunătățirea proprietăților nutriționale și funcționale ale globulinelor de rezervă, precum și pentru a reduce alergenicitatea lor, pot avea succes numai atunci când nu încalcă baza fiziologică a funcționării lor în timpul maturării și germinării semințelor. Soluția acestei probleme ar trebui să se bazeze pe rezultatele studiului privind legitățile de formare a structurilor conservative ale globulinelor 7S și 11S în timpul evoluției lor, precum și pe mecanismele care reglează degradarea globulinelor de rezervă în timpul germinării semințelor, perioada cea mai vulnerabilă în ontogeneza plantelor.

În legătură cu problemele de mai sus, globulinele de rezervă au fost și rămân un obiect important al cercetărilor, realizate de la sfârșitul secolului trecut până în zilele noastre, de cele cinci echipe de frunte conduse de J. M. Dunwell (Marea Britanie), S. Utsumi (Japonia), K. Müntz (Germania), H. Bäumlein (Germania) și I. A. Vaintraub (mai târziu A. D. Șutov, Moldova).

Echipele conduse de H. Bäumlein și I. A. Vaintraub, mai târziu A. D. Șutov au stabilit similitudinea între globulinele de rezervă și germine (oxalat-oxidazele) la nivelul structurilor primare și superioare, indicând originea posibilă a globulinelor de rezervă de la sferuline, proteinele eucariotelor inferioare, înrudite cu germinele [4]. Însă, mai târziu, echipa condusă de J.M. Dunwell, a arătat asemănarea, la nivelul structurilor primare, între globulinele de rezervă și nu numai germine/sferuline, ci și oxalat-decarboxilazele bacteriene [5]. Ulterior, asemănarea între globulinele de rezervă, germine și oxalat-decarboxilaze a fost stabilită nu doar la nivelul structurilor primare, ci și superioare.

Astfel, întrebarea predecesorilor străvechi evolutivi ai globulinelor de rezervă (și, prin urmare, a etapelor cheie ale evoluției lor) a rămas deschisă [6]. Prima parte a lucrării prezente este dedicată soluției acestei probleme.

A fost stabilită structură terțiară bidomenică a leguminelor, globulinele de rezervă 11S din semințe. Fiecare dintre domeniile N- și C-terminale sunt formate dintr-un  $\beta$ -barrel conectat la un grup de  $\alpha$ -helixuri [7, 8]. Este prezentată similaritatea structurilor domeniilor N-terminale ( $\alpha$ catenele) și a domeniilor C-terminale ( $\beta$ -catenele). S-a determinat prezența în subunitățile leguminelor a trei regiuni nestructurate extinse – în regiunea centrală a  $\beta$ -barrel-ului și între  $\beta$ barrel și  $\alpha$ -helixuri în  $\alpha$ -catene, precum și în linkerul extins, care leagă  $\alpha$ - și  $\beta$ -catenele.

Echipa lui I. A. Vaintraub, mai târziu echipa lui A. D. Șutov (ulterior echipa lui K. Müntz) au stabilit sensibilitatea crescută la proteoliza limitată a globulinelor 11S a tuturor celor trei regiuni enumerate mai sus. S-a arătat ulterior că, proteoliza leguminei din soia sub acțiunea papainei începe cu scindarea linkerului interdomenic și se termină cu scindarea regiunii între  $\beta$ -barrel și  $\alpha$ helixuri [9]. Regiunea clivabilă a  $\alpha$ -helixurilor ale  $\alpha$ -catenelor și a linkerului interdomenic este hidrolizată până la peptide cu masa moleculară mică, în timp ce  $\beta$ -catenele rămân intacte. La această etapă, proteoliza limitată este finalizată, iar scindarea profundă ulterioară a moleculei de legumină, modificată prin proteoliză limitată, are loc prin mecanismul proteolizei "una câte una".

Pe baza analizei cineticii proteolizei, s-a demonstrat că molecula de legumină nativă din soia este insensibilă la proteoliza "una câte una". Această sensibilitate la proteoliza "una câte una" este dobândită datorită modificărilor conformaționale, rezultate în urma proteolizei limitate. Aceste modificări și informează molecula de legumină despre abilitatea de degradare profundă prin mecanismul "una câte una". Rezultate similare au fost obținute la studiul cineticii proteolizei leguminei din semințele de dovleac cu papaină [10]. Astfel, s-a stabilit experimental că proteoliza limitată are rol de reglare în timpul mobilizării proteolitice a leguminelor din soia și din dovleac.

Proteoliza limitată are oare rol de reglare în mobilizarea leguminei din semințele altor plante? Pentru a răspunde la această întrebare, în a doua parte a acestei lucrări, am studiat mecanismul degradării proteolitice a leguminei din ginkgo, unul dintre reprezentanți străvechi ai spermatofitelor. În cazul unui răspuns pozitiv la această întrebare, se poate presupune că proteoliza limitată are rol de reglare în timpul mobilizarii proteolitice a leguminelor din semințele plantelor care ocupă o poziție intermediară pe ramura evolutivă a spermatofitelor între gimnospermele vechi (de exemplu, ginkgo) și angiospermele foarte dezvoltate (de exemplu, soia și dovleac).

Vicilinele, globulinele de rezervă 7S bidomenice din semințe, sunt omoloage cu leguminele la nivelul secvențelor de aminoacizi și structurilor terțiare [1, 2]. Cu toate acestea, vicilinele diferă de legumine după mai multe trăsături. Astfel, structura primară a vicilinelor, grupului convicilinelor, conține o extensie N-terminală nestructurată în domeniului N-terminal, care lipsește în legumine. Spre deosebire de legumine, secvența de aminoacizi din domeniul N-terminal al vicilinelor nu conține regiuni nestructurate extinse. O astfel de regiune nestructurată este caracteristică regiunii centrale a  $\beta$ -barrel-ului din domeniul C-terminal al vicilinelor. În cele din urmă, structurile cuaternare ale vicilinelor (trimerică) și leguminelor (hexamerică) sunt, de asemenea, diferite.

Au fost obținute informații cu privire la legitățile proteolizei limitate a vicilinelor (echipele lui I. A. Vaintraub, mai târziu A. D. Șutov și K. Müntz) [11, 12, 13, 14, 15]. S-a stabilit sensibilitatea la proteoliza limitată a celor două regiuni nestructurate menționate mai sus (extensia N-terminală și bucla din partea centrală a  $\beta$ -barrel-ului din domeniul C-terminal), precum și a linkerului interdomenic nestructurat. Este prezentată hipersensibilitatea la proteoliza limitată a tuturor celor trei regiuni enumerate. Pe baza analizei cineticii proteolizei, s-a constatat că mobilizarea vicilinelor are loc ca urmare a proceselor limitate și "una câte una", care au loc simultan independent unul față de celălalt [14].

Până în prezent, literatura științifică mondială a acumulat informații vaste despre legitățile degradării proteolitice a globulinelor de rezervă din semințele multor plante, inclusiv a celor utilizate în alimentație [2]. Rezultatele obținute indică sensibilitate la proteoliza limitată a regiunilor menționate mai sus în structura leguminelor și a vicilinelor. Unele dintre aceste proteine (în principal din arahide și din soia) sunt alergeni (conform SDAP) [3]. Cel puțin o parte din sectoarele secvenței globulinelor de rezervă, care determină alergenicitatea lor, aparțin straturilor de suprafață, eliminate în timpul proteolizei limitate. De exemplu, prelungirea N-terminală a

globulinei 7S din arahide, care este potențial sensibilă la proteoliza limitată și determină, în mare măsură, alergenicitatea acestei proteine [16]. În mod similar, sensibilitatea crescută la proteoliza limitată a regiunii C-terminale din domeniul N-terminal a subunităților globulinei 11S din soia a fost stabilită experimental. Această regiune determină complet alergenicitatea subunității Gly m G1 a globulinei 11S din soia [17].

Exemplele de mai sus indică posibilitatea potențială de reducere a alergenicității globulinelor de rezerva 7S și 11S din arahide și din soia (și, eventual, a altor plante) prin proteoliză limitată. Dacă această ipoteză este adevărată, proteoliza limitată poate fi utilizată ca o modalitate de a reduce alergenicitatea globulinelor de rezervă din semințe. A treia parte a prezentei lucrări este dedicată răspunsului la această întrebare.

**Scopul studiului** constă în studierea legităților formării structurii conservative a globulinelor de rezervă din semințe în timpul evoluției, stabilirea mecanismului degradării proteolitice profunde a acestora și analiza perspectivelor de utilizare a proteolizei limitate a globulinelor de rezervă pentru a reduce nivelul de alergenicitate a acestora.

#### **Obiectivele studiului** sunt:

- analiza relațiilor filogenetice dintre globulinele de rezervă din semințe și proteinele structural omoloage ale procariotelor și eucariotelor;
- stabilirea etapelor succesive în formarea structurii globulinelor de rezervă, care determină funcția lor în semințe;
- studierea cineticii proteolizei globulinei de rezervă 11S din semințele de ginkgo Ginkgo biloba unul dintre cei mai vechi reprezentanți ai spermatofitelor și descrierea mecanismului de reglare a degradării lor proteolitice profunde;
- studierea reacțiilor succesive a proteolizei limitate ale globulinelor 11S și 7S din arahide *Arachis hypogaea*, semințele cărora au un nivel ridicat de alergenicitate;
- stabilirea modificărilor structurale în globulinele 11S și 7S în timpul proteolizei și analiza posibilității de reducere a alergenicității globulinelor de rezervă din arahide prin proteoliza limitată.

#### Ipoteza de cercetare:

1. Evoluția globulinelor de rezervă și a germinelor/sferulinelor s-a dezvoltat după două scenarii independente. Precursorii evolutivi ai globulinelor de rezervă sunt oxalat-decarboxilazele bacteriene. Precursorii evolutivi vegetali ai globulinelor de rezervă din semințe sunt proteinele algelor verzi, structural înrudite cu oxalat-decarboxilazele bacteriene.

2. Proteoliza limitată a leguminei din ginkgo inițiază degradarea profundă a acesteia prin mecanismul "una câte una". Mecanismul de reglare a degradării proteolitice masive a globulinelor de rezervă, care a apărut la un nivel atât de timpuriu de dezvoltare a structurilor primare și superioare ale globulinelor de rezervă ca Gingkophyta, s-a păstrat în timpul evoluției ulterioare a structurilor globulinelor din familia leguminelor.

3. Proteoliza limitată a leguminei și a vicilinei din arahide duce la eliminarea sectoarelor secvențelor de aminoacizi în care sunt identificați determinanții antigenici (epitopii IgE) care determină alergenicitatea acestor proteine. Proteoliza limitată poate fi recomandată ca o modalitate de a reduce nivelul de alergenicitate a globulinelor de rezervă din semințele de arahide și soia, precum și a unui număr de alte plante utilizate în alimentație.

Sinteza metodologiei de cercetare. În conformitate cu scopul și obiectivele trasate, în vederea demonstrării ipotezelor de cercetare înaintate, au fost aplicate metode de cercetare clasice și moderne.

La analiza structurilor primare și superioare ale proteinelor au fost utilizate metodele moderne, furnizate de următoarele programe internaționale: BLAST (<u>http://www.ncbi.nih.gov/</u>) – pentru a căuta secvențele omoloage de aminoacizi ale proteinelor; Clustal W2 și Clustal Omega – pentru alinierea (alignment) secvențelor de aminoacizi găsite ale proteinelor; TREECON – pentru analiza evolutivă a proteinelor găsite; DeepView Swiss-PDB Viewer – pentru analiza structurilor terțiare ale proteinelor, construirea diagramelor panglică și alinierea structurală; Swiss Model – pentru a construi structurile terțiare și cuaternare model ale proteinelor studiate; programul <u>http://cib.cf.ocha.ac.jp/bitool/ASA/</u> a fost utilizat pentru a determina suprafața resturilor de aminoacizi accesibilă la solvent în structurile de cristal și model ale oligomerilor globulinelor 11S.

Extragerea și purificarea proteinelor studiate a fost efectuată prin fracționarea cu sulfat de amoniu și cromatografia hidrofobă pe coloana cu phenil-Sepharose CL-4B. Determinarea purității preparatelor proteice obținute, precum și identificarea produselor de proteoliză a fost realizată prin SDS-electroforeză.

Fiabilitatea rezultatelor obținute este asigurată de utilizarea metodelor statistice standard pentru evaluarea rezultatelor analizelor evolutive (pragramul TREECON), folosirea metodei computerizate moderne pentru cuantificarea rezultatelor electroforezei (programul Phoretix 1D Gel Analysis v.5.10), precum și de absența contradicțiilor dintre rezultatele proprii și datele publicate.

#### Sumarul capitolelor tezei.

Teza include: adnotare prezentată în limba română, engleză și rusă, lista tabelelor, lista figurilor, lista abrevierilor, introducere, patru capitole, primul dintre care prezintă analia situației din domeniu, iar următoarele trei reflectă contribuțiile metodice și rezultatele propriilor cercetări, concluzii generale și recomandări, bibliografie, declarația privind asumarea răspunderii și CV-ul autorului.

În **Introducere** este argumentată actualitatea problemei și importanța temei abordate, sunt formulate scopul și obiectivele cercetării, se prezintă ipotezele de cercetare, sinteza metodologiei de cercetare și justificarea metodelor de cercetare alese, precum și sumarul capitolelor tezei.

**Capitolul 1 "Globulinele de rezervă din semințe"** conține o analiză a datelor din literatura de specialitate despre structurile primare și superioare ale globulinelor de rezervă din semințe, despre mecanismele de proteoliză a acestora și despre reglarea acestui proces; oferă informații despre alergenicitatea globulinelor de rezervă din semințele de arahide, soia și a altor plante.

**Capitolul 2 "Originea și evoluția structurilor globulinelor de rezervă din semințe"** include metodele și strategia de cercetare, precum și descrierea căilor independente de evoluție a globulinelor de rezervă din semințe și a germinelor din plante și a sferulinelor din eucariote inferioare, înrudite structural cu globulinele de rezervă.

**Capitolul 3 "Legitățile proteolizei globulinei 11S din semințele de ginkgo"** include obiectul de studiu și metodele de cercetate și reflectă rezultatele privind proteoliza limitată a globulinei 11S din semințele de *Ginkgo biloba*. S-a stabilit rolul proteolizei limitate a globulinei 11S din semințele de ginkgo în inițierea degradării sale proteolitice profunde și a fost propusă ipoteza privind existența unui mecanism de reglare a acestui proces, caracteristic întregii familii de globuline 11S.

**Capitolul 4 "Proteoliza ca modalitate de reducere a alergenicității globulinelor de rezervă din arahide"** cuprinde obiectul de studiu și metodele de cercetare, precum și rezultatele studierii proteolizei globulinelor de rezervă 11S și 7S din semințele de arahide. Sunt descrise legitățile proteolizei globulinelor de rezervă din semințele de arahide și sunt analizate perspectivele reducerii alergenicității globulinei de rezervă din arahide și soia prin metoda proteolizei limitate.

# 1. GLOBULINELE DE REZERVĂ DIN SEMINȚE

Termenii viciline și legumine au fost introduși la sfârșitul secolului XIX de către T. B. Osborne pentru a desemna două tipuri de proteine de rezervă din semințele a patru leguminoase [18]. În prezent s-a stabilit, că vicilinele și leguminele formează două familii conservative de proteine omoloage din semințele tuturor spermatofitelor [1, 2, 19-21]. Indiferent de diferențele în proprietățile globulinelor de rezervă, în cadrul fiecăreia dintre aceste familii este obișnuit să le numim în funcție de valorile aproximative ale coeficienților de sedimentare [22, 23] globuline 7S (viciline) și globuline 11S (legumine).

#### 1.1. Structura globulinelor de rezervă din semințe

**Globulinele 7S.** Structura primară a globulinelor de rezervă ale acestei familii poate fi ilustrată pe exemplul  $\beta$ -conglicininei, globulina 7S din soia *Glycine max* [24, 25]. Molecula trimerică a  $\beta$ -conglicininei este formată dintr-o combinație aleatorie a subunităților de trei tipuri –  $\alpha$ ,  $\alpha'$  și  $\beta$  [26]. În fiecare dintre aceste subunități există domeniile N- și C-terminale structural echivalente. Structura fiecăruia dintre aceste domenii (Figura 1.1) este formată dintr-un  $\beta$ -barrel, cu  $\beta$ -strandurile antiparalele *BCDEFGH*, unit cu grupul de  $\alpha$ -helixuri *h1*, *h2* și *h3*;  $\alpha$ -helixul *h4* suplimentar este caracteristic numai pentru domeniul N-terminal. Structura  $\beta$ -barrel-ului domeniilor N- și C-terminal este suplimentată cu  $\beta$ -strandurile *A*, *A'* și *J*, *J'* și  $\alpha$ -helixul *h0*, precum și cu  $\beta$ -strandul *Z*, care stabilizează interacțiunile intersubunitare în trimer.

În subunitatea  $\alpha'$  a globulinei 7S din soia există trei regiuni nestructurate extinse: a – extinderea N-terminală a domeniului N-terminal și b – bucla între  $\alpha$ -helixul h3 și  $\beta$ -strendul J' (Figura 1.2); domeniul C-terminal, bucla între  $\beta$ -strendurile E și F (c). A patra regiune nestructurată d – secvența C-terminală a domeniului C-terminal (Figura 1.2).

Structurile terțiare cunoscute ale globulinelor 7S foarte conservative sunt apropiate (Tabelul 1.1). Excepție face fazeolina, globulina 7S din fasole *Phaseolus vulgaris*, la care în domeniul C-terminal lipsește  $\beta$ -strandul *Z*, iar bucla dintre  $\beta$ -strandurile *E* și *F* este neobișnuit de scurtă. Dimpotrivă, bucla dintre  $\beta$ -strandurile *F* și *G* în domeniul N-terminal este neobișnuit de alungită. Aceste particularități ale structurii fazeolinei sunt discutate mai jos (subcapitolul 1.3.) în legătură cu caracteristicile individuale ale proteolizei acestei proteine.





Liniile punctate corespund regiunilor nestructurate: extinderea N-terminală extinsă și bucla dintre  $\alpha$ -helixul h3 și  $\beta$ -strendul J'.

136	esqreprrhkr	1kNP <u>FHF</u> NSK	REQTLEKNO	<u>YGHVRVLQ</u> RF1	VKRSQQLQNLR
	а	Z	A'	А	
186	DY <u>RILEFNSK</u> I <i>B</i>	NTLLLPHHA C	ADYLIVILI D	N <u>GTAILTLVN</u> N <i>E</i>	ND <u>DRDSYNLQS</u> F
236	GDALRVPAGT	YYVVNPDND: H	EN <u>LRMITLA:</u> I	<u>IPV</u> NKPGRFES	S <u>FFLSST</u> QAQQ J
286	SY <u>LQ</u> GFS <u>KNII</u> h1 h2	<u>leasy</u> dtk <u>fe</u> ?	EINKVLFGre b3	$egin{array}{c} egin{array}{c} egin{array}$	lqes <u>vivei</u> sk <i>J</i> ′
336	KQIRELSKHAN h4	SSSRKTISS	EDKPFNLRSI Z	RDPIYSNKLGH A'	KLFEITPEKNP A
386	QLRDL <u>DVFLS\</u> B	<u>/VDMNEGALF</u>	<u>'LPHFNS</u> KA <u>IV</u> C	<u>/vlvin</u> eg <u>ean</u> D	<u>NIELVGI</u> keqq <i>E</i>
436	qrqqqeeqPL <u>F</u> c	<u>VRKYRAEL</u> S <i>F</i>	EQ <u>DIFVI</u> PAC G	GY <u>PVVVNAT</u> SI <i>H</i>	O <u>LNFFAFGIN</u> A <i>I</i>
486	ENN <u>QRNFL</u> AGS J	KDNV <u>IS</u> QIP <i>h1</i>	SQVQELAFPO h2	GS <u>AKDIENLII</u> <i>h3</i>	sqses <u>yfvda</u> J'
536	$\operatorname{qpqqkeegnkq}_d$	grkgplssil	.ra		



pdb	Subunitatea	Cristalul	Specia	Link
1uik	α′	ABC		[25]
1ipj	β	ABC	Chains man	[24]
1 ipk	β	ABC	Glycine max	[24]
1uij	β	ABCDEF		[27]
2phl	β	ABC	Phaseolus vulgaris	[29]
3s7e		AB*	Arachis hypogaea	[30]
3s7i		AB*		[30]
3smh		ABCDEF		[31]
2cv6		ABC	Vigna radiata	[32]
2ea7		ABC	Viene a secondaria	[33]
2eaa		ABC	vigna angularis	[33]
1cau		А		[28]
2cau		А	Canavalla ensijormis	[34]
4lej		А	Pinus koraiensis	[35]

Tabelul 1.1. Globulinele 7S cu structura terțiară cunoscută.

Notă: \* – subunitățile A și B, ce aparțin trimerilor diferiți în structurile hexamerice. Pentru canavalină, globulina 7S *Canavalia ensiformis*, s-au stabilit variații ale structurii terțiare ale domeniilor N- și C-terminale separat [28] (1caw, 1cax, 1dgr, 1dgw).



**Fig. 1.3. Structura cuaternară a globulinei 7S din soia 1ipk** [24] a, b, c – subunitățile globulinei; N și C – domeniile N- și C-terminale, respectiv.

Subunitățile *a, b, c* în moleculele globulinelor 7S sunt combinate în trimer datorită interacțiunilor dintre regiunile C-terminale ale domeniilor N-terminal și C-terminal ale fiecăreia dintre cele trei perechi de subunități adiacente în trimerul subunităților [24, 25] (Figura 1.3).

În funcție de condițiile de cristalizare, trimerii subunităților ABC ale unor globuline 7S (inclusiv β-conglicinina [27]) formează hexamere 9S ABCDEF (Tabelul 1.1).

**Globulinele 11S**. Structura primară a globulinelor de rezervă din această familie poate fi ilustrată pe exemplul glicininei, globulină 11S din soia *Glycine max* [7, 8, 36]. Subunitățile glicininei sunt sintetizate într-un lanț polipiptidic unic, post-translațional scindate în domeniile Nși C-terminal (respectiv  $\alpha$ - și  $\beta$ -catene) [26]. Molecula hexamerică a glicininei este formată prin combinarea aleatoare a cinci subunități de două tipuri: de tip I: G1 (A1aB1b), G2 (A2B1a), G3 (A1bB1a) și de tip II: G4 (A5A4B3), G5 (A3B4) [26]. În fiecare dintre aceste subunități există domeniile N- și C-terminal structural echivalente. Structura fiecăruia dintre aceste domenii (Figura 1.4 și 1.5) este format din  $\beta$ -barrelul cu  $\beta$ -strendurile antiparalele *BCDEFGHI*, unit cu un grup de  $\alpha$ -helixuri *h1, h2* și *h3*. Structura  $\beta$ -barrelui este suplimentată cu  $\beta$ -strendurile *A, A'* și *J, J'* și  $\alpha$ helixul *h0*. În domeniul C-terminal este prezent  $\beta$ -strendul *Z* suplimentar, care stabilizează interacțiunile interdomeniu în fiecare dintre cele două semimolecule trimerice ale glicininei combinate în hexamer.



#### Fig. 1.4. Diagrama panglică a structurii terțiare a domeniului N-terminal al subunității A3B4 al globulinei 11S din soia pdb|1od5 [8]

Liniile punctate corespund regiunilor nestructurate: N-terminal scurt (*a*), buclei între  $\beta$ -strendurile *E'* și *F'* (*b*), buclei între  $\beta$ -strendul *J* și  $\alpha$ -helixul *h1* (*c*) și linkerul interdomenic (*d*).



Fig. 1.5. Secvența de aminoacizi a subunității A3B4 a globilinei 11S din soia, glicinina pdb|1od5 [8]

Poziția structurilor secundare (β-strendurile *Z-J'* și α-helixurile *h0-h3*) este indicată prin subliniere. Literele mici corespund regiunilor nestructurate *a-d*. Săgeata arată punctul de scindare al secvenței A3B4 în α- și β-catene (domeniile N- și C-terminale, respectiv).

În subunitățile glicininei sunt prezente cinci regiuni nestructurate: a - N- și C-terminal scurte și trei regiuni nestructurate extinse în  $\alpha$ -catene: b – bucla între  $\beta$ -strandurile E și F, c – bucla între  $\beta$ -barrel și  $\alpha$ -helixuri și d – linkerul interdomenic.

Structurile terțiare ale globulinelor 11S foarte conservative, prezentate în Tabelul 1.2, diferă foarte puțin de structura globulinei 11S din soia pdb|1od5 prezentată în Figura 1.4 și 1.5. O trăsătură distinctivă a globulinei 11S din arahide  $3c_3v$  [37] este prezența în domeniul său N-terminal a  $\beta$ -strandului *Z*, care stabilizează interacțiunile interdomeniu. Evidențiem, că  $\beta$ -strandul *Z* este caracteristic atât pentru domeniile N-terminale, cât și pentru cele C-terminale ale globulinelor 7S (Figura 1.2).

La fel ca și în cazul globulinelor 7S, subunitățile *a, b* și *c* sintetizate într-un lanț polipeptidic unic în molecula globulinei 11S sunt inițial combinate în trimer datorită interacțiunilor dintre regiunile C-terminale ale domeniilor N-terminale și C-terminale ale fiecăreia dintre cele trei perechi de subunități adiacente în trimerul subunităților [7] (Figura 1.6). După procesingul proteolitic (scindarea subunităților în  $\alpha$ - și  $\beta$ -catene), doi trimeri în molecula matură a globulinelor 11S se combină într-un hexamer [8, 38, 39].

pdb	Subunitatea	Cristalul	Specia	Link
1fxz	A1aB1b	ABC	Glycine max	[7]
1ucx	1aB1b′	ABC	Glycine max	[36]
1ud1	AlaBlb''	ABC	Glycine max	[36]
10d5	A3B4	AB*	Glycine max	[8]
2dsf	A3B4	AB*	Glycine max	-
2dsh	A3B4	ABCDEF	Glycine max	-
3ksc		ABCDEF	Pisum sativum	[38]
3ehk		ABCDEF	Prunus dulcis	[38]
3fz3		ABCDEF	Prunus dulcis	[39]
3qac		А	Amaranthus hypochondriacus	[40]
2evx		А	Cucurbita maxima	-
2e9q		А	Cucurbita maxima	[38]
3kgl		ABCDEF	Brassica napus	[38]
3c3v		A	Arachis hypogaea	[37]

Tabelul 1.2. Globulinele 11S cu structura terțiară cunoscută.

Notă: \* – subunitățile A și B, ce aparțin trimerilor diferiți în structurile hexamerice.



**Fig. 1.6. Structura cuaternară a globulinei 11S din soia 1fxz** [24] a, b, c – subunitățile globulinei; N și C – domeniile N- și C-terminale, respectiv.

#### 1.2. Originea și evoluția globulinelor de rezervă

Pentru prima dată informații sigure despre rudenia evolutivă a globulinelor 7S și 11S au apărut pe baza comparării unui număr limitat de secvențe de aminoacizi cunoscute ale domeniilor C-terminale conservative [41, 42]. Mai târziu au fost stabilite structurile cristaline ale globulinei 7S din canavalie *Canavalia ensiformis* [34] și din fasole *Phaseolus vulgaris* [29] și pe baza comparării structurilor primare ale unui număr mare de globuline 7S și 11S a fost arătată similitudinea structurală a subunităților din proteinele acestor familii, cu mare probabilitate, indicând rudenia evolutivă a acestora [29]. Cu toate acestea, doar mai târziu a fost obținută dovada absolută a rudeniei evolutive a globulinelor 7S și 11S, bazată pe coincidența stabilită a structurii exon-intron a genelor lor [43].

După apariția informațiilor despre structura cristalină pdb|1fxz a globulinei 11S din soia [7], rudenia presupusă a globulinelor 7S și 11S a fost confirmată pe deplin. Astfel, la suprapunerea structurilor terțiare ale globulinei 7S din canavalie pdb|2cau și a globulinei 11S din soia pdb|1fxz, abaterea standard a poziției 191 atomi C<sup> $\alpha$ </sup> în structura scheletului acestor proteine nu depășește 1,3 Å [7]. Rudenia structurală a globulinelor 7S și 11S este confirmată de suprapunerea parțială sau chiar completă a poziției structurilor secundare din secvențele acestor proteine (Figura 1.7 și 1.8).

Domeniile N- și C-terminale ale globulinelor 7S și 11S sunt echivalente din punct de vedere structural [43]. În plus, secvența tuturor celor patru domenii în aceste două proteine sunt omoloage în toate combinațiile de perechi posibile, ceea ce indică originea lor de la un predecesor monodomenic comun, supus duplicării [43]. Un astfel de predecesor poate fi germina, proteină vegetală monodomenică, a cărei secvență de aminoacizi este omoloaga cu secvențele atât a domeniului N-terminal, cât si C-terminal ale globulinelor de rezervă din seminte (ipoteza 1) [44].

S-a stabilit că structura cristalină a germinei pdb|1fi2 din orz *Hordeum vulgare*, la fel ca și în cazul globulinelor de rezervă, este constiuită din  $\beta$ -barrel cu  $\beta$ -strandurile antiparalele legate cu un grup  $\alpha$ -helixuri [45]; în plus, se constată o suprapunere parțială sau chiar completă a poziției structurilor secundare în regiunea  $\beta$ -barrelui; abaterea standard a poziției 100 atomi C<sup> $\alpha$ </sup> în structura scheletului acestor proteine nu depășește 1,1 Å (Figura 1.9 și 1.10).



Fig. 1.7. Diagramele panglică ale structurilor domeniilor C-terminale ale globulinei 11S din soia 1fxz (a) și globulinei 7S *Canavalia ensiformis* 2cau (b) suprapuse spațial, dar arătate separat [7]

Linia întreruptă corespunde regiunii nestructurate între  $\beta$ -strandurile *E* și *F*, caracteristică globulinelor 7S.

		a Z A' A h0
1FXZ	1	fssreqpqqNECQIQKLNALKPDNRIESEGGLIETWNPN-NKPFQCA
2CAU	41	esRAQNNP <u>YLF</u> RSN <u>KFLTLFKNQHGSLRLLQ</u> RFNEDTEKLENLR
		.: . : *: * :
		B C D E
1FXZ	47	GVALSRCTLNRNALRRPSYTNGPQEIYIQQGKGIFGMIYPGCPSTfeepq
2CAU	85	DYRVLEYCSKPNTLLLPHHSDSDLLVLVLEGQAILVLVNPD
		F G H I
1FXZ	97	qpqqrgqssrpqdRHQKIYNFREGDLIAVPTGVAWWMYNNE-DTPVVAVS
2CAU	126	GRDTYKLDQGDAIKIQAGTPFYLINPDNNQNLRILK
		· * : : : · * * · * · : · · · · · · · · · · · · · · · · ·
		л b
1FXZ	146	$\underline{\texttt{IIDT}\texttt{NSLENQLDQ}\texttt{MPRR}}\underline{\texttt{FYL}} \\ \texttt{AGNQEQEFLKYQQeqgghqsqkgkhqqeee}$
2CAU	162	FAITFRRPGTVEDFFLSSTKRL
		. * . **.:
		h1 h2 h3 J'
1FXZ	196	neGGSILSGFTLEFLEHAFSVDKQIAKNLQGEnegedKGAIVTVKGGLSV
2CAU	184	PSYLSAFSKNFLEASYDSPYDEIEQTLLQEEQEGVIVKMPkdqiq
		* **.*: :*** ::. : : : : :*.**::
1FXZ	246	${\tt IKPptdeqqqrpqeeeeeedekpqckgkdkhcqrprgsqsksrrngide}$
2CAU	229	eiskhaqsssrktl
		Z A' A h0 B
1FXZ	296	tICTMRLRHNIGQTSSPDIYNPQAGSVTTATSLDFPALSWLRLSAEFGSL
2CAU	243	SSQDKPFNLRSRDPIYSNN-YGKLYEITPEKNSQLRDLDILLNCLQM
		* *.: * . * * : : .:
		C D E C F
1FXZ	346	RKNAMFVPHYNLNANSIIYALNGRALIQVVNCNGERVFD
2CAU	289	NEGALFVPHYNSRATVILVANEGRAEVELVGLEqqqqqglesmQLRR-YA
		.:.*:***** .*. *: * :*** :::*. : * :
		G H I J
1FXZ	386	GELQEGRVLIVPQNFVVAARSQSDNFEYVSFKTNDTPMIGTLAG-ANSL
2CAU	338	<u>ATLSEGDIIVIPSSFPVALKAA</u> SDLNMVGIGVNAENNERNFLAGHKENV
		. *.** ::::** ** :: ** . *** *** :.:
		h1 h2 h3 J' d
1FXZ	433	LNALPEEVIQHTFNLKSQQARQIKNNNPFKFLVPPQESqkrava
2CAU	387	IRQIPROVSDLTFPGSGEEVEELLENQKESYFVDGQPRHidaggkarrah
		·· ·* ··* · ** ···· ··· ** ····* *

#### Fig. 1.8. Alinierea structurală a secvențelor de aminoacizi ale globulinei 118 *Glycine max* 1fxz și globulinei 78 *Canavalia ensiformis* 2cau [7]

Regiunile nestructurate sunt arătate cu litere mici. Poziția structurilor secundare ( $\beta$ -strendurile *Z-J'* și  $\alpha$ -helixurile *h0-h3*) este indicată prin subliniere. Sunt prezentate resturile de aminoacizi identice (\*), substituțiile conservative (:) și indiferente (.).



Fig. 1.9. Diagramele panglică ale structurii terțiare ale domeniului N-terminal a globulinei 78 *Canavalia ensiformis* 2cau (a) și germinei pdb|1fi2 din orz (b) suprapuse spațial, dar prezentate separat [45]

Z A' А h0 в C 39 SSKLTKAGNTSTPNGSAVTELDVAEWPGTNTLGVSMNRVDFAPGGTNPP 1fi2 2cau 248 PFNLRSRDPIYSNNYGKLYEITPEKNSQLRDLDILLNCLQMNEGALFVP : :\* . . : \* . : \*: : \*.: :\* ::: D Е F 1fi2 88 HIHPRATEIGMVMKGELLVGILGSLDS-----GNKLYSRVVRAGET 2cau 297 HYNSRATVILVANEGRAEVELVGLEQQQQQGLESMQLRRYAATLSEGDI : \*\*\* \* :. :\*. \* ::\* :. : \*: .: \*: J G н h1 Т 1fi2 129 FVIPRGLMHFQFNVGKTEAYMVVSFNSQNPGIVFVPLTLFGSDPPIPTP 2cau 346 IVIPSSFPVAL-KAASDLNMVGIGVNAENNERNFLAGHKENVIRQIPRQ :\*\*\* .: :... : :..\*::\* \* : J' h2 h3 1fi2 178 VLTKALRVE-AGVVELLKSK----FAGGS 2cau 394 VSDLTFPGSGEEVEELLENQKESYFVDGQ :: \*\*\* • \* \* \*

Fig. 1.10. Alinierea secvențelor de aminoacizi ale domeniului N-terminal al globulinei 7S *Canavalia ensiformis* 2cau și al zonei modulului structural β-barrel-α-helixuri al germinei 1fi2 [45]

Poziția structurilor secundare ( $\beta$ -strendurile *Z-J'* și  $\alpha$ -helixurile *h0-h3*) este indicată prin subliniere. Sunt prezentate resturile de aminoacizi identice (\*), substituțiile conservative (:) și indiferente (.).

Ipoteză 1 despre înrudirea evolutivă a germinelor și globulinelor de rezervă din semințe a stimulat căutarea ulterioară a predecesorilor evolutivi ai acestor proteine, ceea ce a dus la descoperirea unei vaste superfamilii de proteine înrudite structural numite cupine [5, 46-49]. Căutarea s-a bazat pe două motive nodale ale secvențelor de aminoacizi, din structura germinei [49], care corespund la două jumătăți scurte ale β-barrelui: Gx5HxHx3,4Ex6G (β-strendurile *CDE*) și Gx5PxGx2Hx3N (β-strendurile *FGH*).

Băncile de date moderne conțin multe mii de secvențe de aminoacizi, atribuite superfamiliei cupinelor (clan cl09118). Prezența reală a  $\beta$ -barrelui, care este considerată a fi caracteristică pentru toate cupinele, a fost arătată în structurile terțiare a mai mult de o sută de proteine din arhibacterii, bacterii și eucariote. Se presupune că cel puțin 18 subclase de cupine, provin dintr-un predecesor evolutiv comun [49]. Cu toate acestea, omologia regiunilor scurte analizate ale secvențelor de aminoacizi ale cupinelor între subclasele menționate nu este întotdeauna evidentă. Asemănarea observată între structurile primare ale cupinelor în multe cazuri poate avea un caracter convergent din cauza formării unor structuri terțiare similare [2].

Din punct de vedere al evoluției globulinelor de rezervă, dintre celelate cupine un interes deosebit prezintă familia oxalat-decarboxilazelor bacteriene (OD). Structura OD este formată din două domenii omoloage domeniilor globulinelor de rezervă atât în regiunea  $\beta$ -barrelui, cât și în regiunea  $\alpha$ -helixurilor [2] (Figura 1.11 și 1.12).



Fig. 1.11. Diagramele panglică ale structurilor terțiare ale domeniilor C-terminale ale globulinei 118 din soia *Glycine max* pdb|lod5 (a) și OD cianobacteriilor *Synechocystis sp.* pdb|2vqa (b) suprapuse spațial, dar prezentate separat [2]

lod5	14	LNAL-EPDHRVE	SEGGLIET	WNSQH-P	ELQCAGVTV:	SKRTLNRNGL	LPSYSPYPQMI
2vqa	21	FTYAFSKTPLVL	YD <b>GG</b> TTK(	OVGTYNF P	VSKGMAGVYI	ISLEPG-AIRE	LHWHANAAEWA
		<u>Z</u> A	' A		h0 1	в (	<u> </u>
lod5	70	I <b>V</b> VQ <b>G</b> KGAIGFA	F <b>P</b> G[25]S	SHQKI RHF	NEGDVLVIPI	GVPYWTYNT	DEPVVAISLLD
2vqa	78	YVMEGRTRITLT	SPE0	GKVEIADV	DKGGLWYFPI	RGWGHSIEGI	PDTAKFLLVFN
		Е	_	F	G	H	I
lod5	149	TSNFNQLNDQNP	RVFYL[36	6]GSV <b>LS</b> G	FSKHFLAQSI	NTNEDTAEKI	RSPDDERKQIV
2vqa	132	DGTFSEGA	-TFSV	TDWLSH	TPIAWVEEN	LGWTAAQVAQI	PK-KQVYISSY
			J	h1	h2	h3	<i>J'</i>
lod5	239	TVEGG[86]KLH	ENIARPSE	RADFYNPK	AGRISTLNSI	LTLPALROFGI	SAQYVVLYRNG
2vqa	180	GPASG[14]EVP	HTHNLLG	QPLVSLG	GNELRLASA	EFPGSFN	TGALIHLEPGA
			Ζ	A'	A	h0	В
lod5	379	IYSPHWNLNANS	VI <b>YV</b> TRGE	KGRVR <b>V</b> VN	COGNAVFDGI	ELRRGQLLVVI	ONFVVAEQGGE
2vqa	246	MRQL <b>HW</b> HPNADE	WQYVLDGH	EMDLTVFA	SEG-KASVSI	RLQQGDVGYVI	KGYGHAIRNSS
		С	D	Е	F	G	H
lod5	437	QGLEYVVFKTHH	NAVSSYI-	KDVFRAI	PSEVLSNSYI	NLGQSQVRQL	YQGNSGPLVNP
2vqa	303	QKPLDIVVVFND	GDYQSIDI	LSTWLASN	<b>PSSVLGN</b> TF(	DISPELTKKL	VQDTIFSLP
		I	J	h1	h2	h3	J'

Fig. 1.12. Alinierea secvenței de aminoacizi ale subunităților globulinei 11S din soia *Glycine* max 10d5 și OD cianobacteriei *Synechocystis sp.* 2vqa [2]

Poziția structurilor secundare ( $\beta$ -strendurile *Z-J'* și a  $\alpha$ -helixurilor *h0-h3*) este indicată prin subliniere. Prin text gras sunt arătate resturile de aminoacizi identice. Numerele din parantezele pătrate indică poziția și numărul de resturi omise în cele trei regiuni nestructurate ale secvenței globulinei 11S.

Aceasta permite de a presupune, că anume oxalat-decarboxilazele bacteriene ar putea fi precursori direcți și unici ai globulinelor de rezervă [49] (ipoteza 2). O dovadă semnificativă în favoarea acestei ipoteze este detectarea proteinei bidomenice oxalat-decarboxilazo-similară din algele verzi ca o legătură intermediară evolutivă între oxalat-decarboxilazele bacteriene și proteinele legumino-/vicilino-similare din struțișor și ferigă [2]. Acestea din urmă sunt, aparent, cei mai apropiați precursori ai globulinelor de rezervă din semințe. Acest lucru, în special, este evidențiat de identitatea structurii exon-intron (coincidența numărului, poziției și fazei intronilor) a genei, care codifică globulinele 7S și 11S și, respectiv proteinele vicilino- și legumino-similare din struțișor [2].

Se pare că intriga originii globulinelor de rezervă din semințe tinde să favorizeze oxalatdecarboxilazele bidomenice bacteriene. Cu toate acestea, în cazul analizei filogenetice, cu utilizarea ca claster rădăcină a unor proteine monodomenice fie ale bacteriilor [2] sau ale arhibacteriilor [6], globulinele de rezervă din semințe și germinele (precum și precursorii lor sferulinele [44]) formează un cluster comun, în mod egal înrudit cu oxalat-decarboxilazele bacteriene (Figura 1.13) [2]. Astfel, problema originii globulinelor de rezervă din semințe, în esență, rămâne deschisă.



#### Fig. 1.13. Analiza filogenetică a secvențelor de aminoacizi ale globulinelor de rezervă (și a precursorilor acestora din plantele cu spori), germinelor (și precursorii lor sferulinelor) și OD [2]

Regiunea de secvență analizată (167 de poziții ale alinierii) corespunde domeniilor C-terminale ale proteinelor bidominice (\*\*) și modulului structural β-barrel-α-helixuri, ale proteinelor monodominice (\*). Numerele de deasupra ramurilor reflectă suportul statistic al clusterelor (% din 1000 de replicări).

#### 1.3. Proteoliza globulinelor de rezervă și reglarea acesteia

Pentru semințele germinative, sunt cel mai mult caracteristic două clase de proteinaze, implicate în degradarea globulinelor de rezervă [50]: proteinazele asemănătoare papainelor cu specificitate de substrat îngustă [50-52] și legumainele Asn-specifice [53-55]. În semințele mature legumainele catalizează scindarea post-translațională a globulinelor 11S în domeniile N- și C-terminale (respectiv,  $\alpha$ - și  $\beta$ -catene) [56-58]. Proteazele aspartice [59-61] și metaloproteinazele [62-64] participă, de asemenea, la degradarea globulinelor de rezervă în semințele germinative.

*Mecanismele proteolizei proteinelor.* Mecanismele proteolizei proteinelor sunt determinate nu atât de specificitatea primară de substrat a proteinazei, cât de caracteristicile individuale ale structurii terțiare ale substratului proteic. Există două mecanisme de proteoliză ale proteinelor – limitată și nelimitată ("una câte una", extinsă) [65-67]. Diferențele principale dintre aceste mecanisme sunt prezentate în Tabelul. 1.3.

Proteoliza limitată	Proteoliza "una câte una" "one-by-one"		
Concentrația în greutate a proteinei se micșorează ca urmare a reducerii masei moleculare a substratului	Concentrația în greutate a proteinei se micșorează ca urmare a unei degradări complete a moleculelor de substrat		
Concentrația molară a proteinei reziduale este constantă în timpul reacției	Masa moleculară a proteinei reziduale este constantă în timpul reacției		
Procesul se termină după epuizarea legăturilor sensibile în moleculele substratului proteic	Procesul se termină după epuizarea moleculelor substratului proteic		
Procesul se desfășoară în ritm descrescător și se termină cu formarea unui produs macromolecular stabil	Proteoliza "una câte una" este o reacție de ordinul pseudo-unu [65, 66]		

Tabelul 1.3. Caracteristicile principale ale reacțiilor idealizate ale proteolizei limitate și "una câte una"a proteinelor native [2].

Proteoliza limitată a proteinei este determinată de prezența în structura sa nativă a legăturilor peptidice cu sensibilitate crescută la atacul proteolitic. Un exemplu tipic de proteoliză limitată este procesingul proteolitic, în care poziția punctului de scindare în structura primară a proteinei depinde numai de structura sa terțiară [38].
Proteoliza "una câte una" [67] este determinată de variabilitatea structurii terțiare a proteinei, datorită fluctuațiilor sale termice. Formele denaturate reversibile ale moleculelor (conformeri) diferă în grade diferite de structura statistică medie, fixată în cristalul proteinei. Amplitudinea modificărilor structurale în unul din acești conformeri este suficientă pentru a dezgoli legătura peptidică, care este pusă la dispoziția atacului proteinazei. După clivarea acestei legături, denaturarea conformerului devine ireversibilă. Scindările ulterioare într-o astfel de moleculă se produc cu o viteză crescătoare, ceea ce duce la degradarea ei explozivă. Astfel, viteza proteolizei "una câte una" este limitată la scindarea primei legături peptide în fiecare moleculă. De aceea, la o concentrație a substratului proteic care corespunde părții inițiale a curbei Michaelis, proteoliza "una câte una" este o reacție de ordinul pseudo-unu [65, 66].

În timpul proteolizei de tip mixt procesele proteolizei "una câte una" și limitată decurg simultan, independent unul față de altul, în cazul când proteina nativă este capabilă de proteoliza "una câte una". Capacitatea proteinei native de proteoliza "una câte una" depinde de caracteristicile individuale ale structurii sale terțiare.

Proteoliza limitată este identificată prin evaluarea vizuală a electroforegramelor proteinei inițiale și a produselor de hidroliză cu masă moleculară mare, disociate cu dodecilsulfat de sodiu, prin micșorarea masei moleculare a zonelor polipeptidelor inițiale și/sau apariția zonelor noi.

La baza analizei cineticii proteolizei, permițând diferențierea proceselor limitate și "una câte una", stau datele cantitative privind variația în timp a masei moleculare a unui produs cu masa moleculară mare, ce nu disociază în condițiile reacției și concentrația de masă a lui în hidrolizat [65].

Reducerea concentrației de masă a proteinei P în timpul de reacție t a proteolizei  $(P^0 \rightarrow P^t)$ poate fi determinată de (Tabelul 1.3): a) scurtarea lanțurile polipeptidice a proteinei inițiale ca urmare a proteolizei limitate, adică reducerea masei sale moleculare  $(M^0 \rightarrow M^t)$  și/sau b) reducerea concentrației molare a proteinei ca rezultat al proteolizei "una câte una"când  $P^t < P^0$  pentru  $M^0 =$  $M^t$ . La descrierea proteolizei de tip mixt, este convenabil să se utilizeze valorile relative ale  $P^t$  și  $M^t$ , exprimate în procente din valorile inițiale  $P^0$  și  $M^0$ . Atunci, la proteoliza limitată,  $P\%^t = M\%^t$ , adică dependența M = f(t) descrie cinetica proteolizei limitate. Dependența P = f(t) descrie cinetica proteolizei "una câte una"în absența proteolizei limitate.

În timpul proteolizei de tip mixt, atunci când ambele procese se desfășoară fie simultan, independent unul față de altul, fie succesiv, contribuția proteolizei limitate la scăderea generală a

concentrației de masă a proteinei în hidrolizat ( $P^0 \rightarrow P^t$ ) este egală cu  $M\%^{0-}M\%^t$ . De aici rezultă că concentrația de masă a proteinei, micșorată doar ca rezultat al proteolizei "una câte una"este egală cu

$$P\%^{t} + (M\%^{0} - M\%^{t})$$
(1)

Scăderea relativă a masei moleculare M a proteinei oligomere în timpul proteolizei limitate  $M^t/M^0$  poate fi calculată pe baza ecuației masei moleculare medii a oligomerului

$$M_n = \sum_i n_i m_i \left(\sum_i n_i\right)^{-1}$$
(2)

(unde  $n_i$  este numărul absolut de monomeri i cu o masa  $m_i$  în compoziția oligomerului). Conform rezultatelor electroforezei, la fiecare etapă a proteolizei sunt estimate cantitativ masa moleculară  $m_i$  a polipeptidei i și conținutul ei relativ de masă  $p_i$ . Deoarece  $n_i = p_i/mi$ , ecuația (2) ia forma

$$M_n = \sum_i p_i (\sum_i p_i/m_i)^{-1},$$
 (3)

*Proteoliza limitată a globulinelor 7S și 11S.* Capacitatea globulinelor de rezervă pentru proteoliza limitată este determinată de prezența în subunitățile lor a regiunilor nestructurate hidrofile expuse la suprafața oligomerilor, menționate mai sus (subcapitolul 1.1).

**Globulinele 7S.** A fost descrisă proteoliză limitată a două globuline 7S, care depinde predominant de caracteristicile individuale ale structurii lor terțiare (Tabelul 1. 4). În timpul hidrolizei *in vivo* [11, 12, 68], ca și în cazul atacului tripsinei și proteinazelor cisteinice C2 [69] și Cpph [11, 70, 71] din semințele germinate de soia și fasole, respectiv, în toate cazurile este arătată scindarea linkerului interdomenic și de regulă, scindarea regiunii N-terminale, extinse în subunitățile  $\alpha$  și  $\alpha'$  ale globulinei 7S din soia (GLYma) și scurtă în alte subunități. Bucla între  $\beta$ strendurile *E* și *F* din partea centrală a  $\beta$ -barrelui domeniului C-terminal, potențial sensibilă la proteoliză, este scindată numai la acțiunea tripsinei.

După cum se va arăta mai jos, un interes deosebit prezintă rezultatele studiului proteolizei limitate a fazeolinei (PHAvu), globulina 7S din fasole, cu legumaina, proteinaza Asn-specifică [11, 12]. Faseolină se deosebește de alte globuline 7S după structura primară și terțiară – linker interdomenic scurt și absența  $\beta$ -strandului Z în domeniul C-terminal [29], precum și a buclei *EF* scurte în acest domeniu și o buclă alungită între  $\beta$ -strendurile F și G în domeniul N-terminal. Aceste caracteristici, precum și specificitatea de substrat îngustă a legumainei determină specificul proteolizei limitate a fazeolinei cu această proteinază: scindarea legăturilor peptidice care implică restul Asn în poziția P1 adiacent regiunii  $\beta$ -strandului A' din domeniul C-terminal, buclei între  $\beta$ - strendurile A și B din același domeniu și buclei FG din domeniul N-terminal (Tabelul 1.4). O serie de date precoce privind legitățile proteolizei globulinelor 7S publicate înainte de stabilirea structurii terțiare ale acestor proteine [50], sunt în concordanță cu datele prezentate în tabelul 1.4.

Tabelul 1.4. Poziția legăturilor peptidice, scindate în timpul proteolizei limitate, în structura terțiară a subunităților globulinei 7S din soia (GLYma) și din fasole (PHAvu).

ZA'ABCDEF GHIJh1-h3J'h4 Z A'A BCDE FGHIJh1-h3J'	Subunitatea	Proteinaza
$\downarrow$ $\downarrow$		in vivo [68]
$\downarrow$ $\downarrow$ $\downarrow$		tripsina [15]
$\downarrow$ $\downarrow$	$GLYma \ \alpha'$	Cpph [11]
$\rightarrow$		C2 [69]
$\rightarrow$		legumaina [11]
$\downarrow$ $\downarrow$		in vivo [68]
$\downarrow$ $\downarrow$	CI Vera er	Cpph [11]
$\rightarrow$	GL Y ma û	C2 [69]
$\rightarrow$		legumaina [11]
$\downarrow$ $\downarrow$		in vivo [11]
$\downarrow$ $\downarrow$		tripsina [14]
$\rightarrow$	CI Vma P	papaina [14]
$\downarrow$ $\downarrow$	GL i ma p	Cpph [11]
$\downarrow$		C2 [69]
$\downarrow$		legumaina [11]
$\downarrow$ $\downarrow$		tripsina [13]
$\downarrow$ $\downarrow$		Cpph [11, 70, 71]
$\downarrow$ $\downarrow$	PHAvu β	C2 [69]
$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$		in vivo [12]
$\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$		legumaina [12, 70]

**Globulinele 11S.** În timpul proteolizei limitate β-catenele (domeniile C-terminale) conservative ale globulinelor 11S rămân intacte [2] (Tabelul 1.5). Proteoliza α-catenelor, fără excepție, ale tuturor subunităților globulinelor 11S începe cu scindarea regiunii C-terminale extinse (linkerul interdomenic) și se termină cu scindarea ulterioară a regiunii C-terminale, care în globulinele 11S native formează α-helixurile h1, h2, h3 și β-strendul J'. Scindarea regiunii α-helixurilor are loc probabil și la proteoliză *in vivo* a subunităților glicininei, globulina 11S din soia [72]. Clivajul buclei alungite între β-strendurile *E* și *F* din domeniul N-terminal depinde de

specificitatea proteinazei [15] și structura primară a globulinei 11S [73, 74], iar la hidroliza homohexamerului glicininei G4 se observă numai la acțiunea prelungită a papainei [15].

### Tabelul 1.5. Poziția legăturilor peptidice, scindate la proteoliză limitată, în structura terțiară a subunităților globulinelor 11S din soia (GLYma G1-G5), din dovleac *Cucurbita maxima* (CUCma), din floarea-soarelui *Helianthus annuus* (HELan), din ovăz *Avena sativa* (AVEsa), din cedru *Pinus sibirica* (PINsi).

A' ABCDE	FGHIJ	h1-h3 3	' A'A	BCDEFGHI J h1-h3 J'	Subunitatea	Proteinaza
		↓	$\downarrow$		GLYma G1	
		$\downarrow$	$\downarrow$		GLYma G2	
		$\downarrow$	$\downarrow$		GLYma G3	papaina [9]
		↓	$\downarrow$		GLYma G4	
		↓	$\downarrow$		GLYma G5	
	↓ ·	↓	$\downarrow$		GLYma G2	tripsina [15]
	↓ ·	↓	$\downarrow$		GLYma G4	papaina [75]
		↓	$\downarrow$		CUCma	papaina [10]
	↓ ·	↓	$\downarrow$		HELan	papaina [73]
	↓ ·	↓	$\downarrow$		AVEsa	papaina [69]
		↓	$\downarrow$		PINsi	papaina [76]

Notă: Regiunea  $\alpha$ -helixuri –  $\beta$ -strendul *J'*, care este scindată la proteoliza limitată, este arătată cu caractere italice.

**Reglarea degradării masive a globulinelor de rezervă.** Procesingul proteolitic al precursorilor proteici este un exemplu al rolului proteolizei limitate site-specifice în formarea structurilor proteinelor mature. Rolul funcțional al procesingului proteolitic este cunoscut (de exemplu, activarea enzimelor).

Proteoliza limitată site-specifică a globulinelor de rezervă, similară procesingului proteolitic se observă în timpul germinării semințelor și *in vitro* [2]. Mobilizarea rapidă a unei regiuni limitate a moleculei globulinelor de rezervă, sensibilă la atacul proteolitic, anticipînd degradarea lor masivă mai lentă, ce decurge în paralel, poate fi cel mai simplu exemplu al rolului fiziologic al proteolizei limitate. Așa cum se arată mai jos, există un alt nivel de reglare mai superior al proteolizei globulinelor de rezervă, stabilit pe baza analizei cineticii a acestui proces.

**Globulinele 7S.** Degradarea masivă *in vivo* a faseolinei este catalizată prin acțiunea combinată a două enzime – proteinaza papaino-similară cu specificitate îngustă Cpph [71] și legumaina Asn-specifică [12]. În timpul hidrolizei faseolinei *in vitro* cu preparatele purificate ale fiecăreia dintre aceste enzime se observă exclusiv proteoliza limitată (Tabelul 1.4), care se încheie cu formarea unui produs macromolecular stabil, insensibil la degradarea prin mecanismul "una câte una" [70, 71]. Numai la acțiunea consecutivă *in vitro* a legumainei și Cpph, așa cum se observă *in vivo*, are loc degradarea completă a fazeolinei. Aceste rezultate demonstrează fără echivoc rolul de reglare a proteolizei limitate în degradarea masivă a fazeolinei prin mecanismul "una câte una". Cu toate acestea, modificările structurale în timpul proteolizei limitate, care comunică fazeolinei capacitatea de a degrada profund, rămân neclare.

În timpul hidrolizei  $\beta$ -conglicininei  $\beta$ 3 din soia (homotrimerul subunităților  $\beta$ ) papaina catalizează nu numai proteoliza limitată, ci și proteoliză "una câte una"[14]. Autorii acestui studiu au concluzionat că proteoliza limitată și "una câte una"a  $\beta$ -conglicininei au loc în paralel independent una față de cealaltă. Într-adevăr, analiza comparativă a accesibilității la solvent a resturilor de aminoacizi din moleculele trimerice native ale globulinei 7S din soia și din fasole arată o densitatea relativ scăzută de împachetare a structurii  $\beta$ -conglicininei, care poate determina sensibilitatea acestei proteine la proteoliza "una câte una"[14].

Totuși, această concluzie nu poate fi considerată suficient justificată din punct de vedere experimental. Autorii au descoperit prezența formelor A și B în preparatul studiat de  $\beta$ conglicinină, din care forma A este scindată complet deja în stadiile inițiale ale reacției. Poate că
forma A corespunde porțiunii denaturate a preparatului inițial de  $\beta$ -conglicinină. Prezența acestei
forme în preparatul de cercetare elimină posibilitatea unei analize cantitative a etapei inițiale a
proteolizei formei native B. Cu toate acestea, numai o astfel de analiză poate oferi un răspuns la
întrebarea sensibilității (sau insensibilității)  $\beta$ -conglicininei native la scindare prin mecanismul
"una câte una".

După cum s-a menționat mai sus, structura fazeolinei este unică din punct de vedere al reglării degradării sale masive. Toate celelalte globuline 7S investigate sunt capabile de degradare profundă (adică, de proteoliza "una câte una") sub acțiunea unei singure proteinaze. Cu toate acestea, rolul proteolizei limitate în scindarea lor profundă rămâne neclar.

**Globulinele 11S**. În timpul hidrolizei cu papaină, moleculele hetero-hexamerice native ale globulinelor 11S din soia [9], dovleac [10] și floarea-soarelui [73], precum și homohexamerul G4

al globulinei 11S din soia [75] sunt insensibile la proteoliza "una câte una". Capacitatea acestor proteine de a degrada masiv prin mecanismul "una câte una"este dobândită numai după finalizarea proteolizei limitate, ceea ce indică rolul său de reglare.

La proteoliza limitată a globulinelor 11S cu papaină este scindată regiunea C-terminală extinsă a  $\alpha$ -catenelor, care cuprinde regiunea  $\alpha$ -helixurilor și  $\beta$ -strendului *J*' (Tabelul 1.5). Se presupune, că sensibilitatea la proteoliza "una câte una", ca consecință a proteolizei limitate, este dobândită în rezultatul modificărilor conformaționale ale globulinei 11S [9].

În timpul hidrolizei globulinei 11S din ovăz cu papaină, proteoliza limitată și "una câte una"decurg paralel, independent una de cealaltă [2], care este determinată de caracteristicile individuale ale structurii acestei proteine (poate densitatea de împachetare relativ scăzută a moleculei sale).

#### 1.4. Globulinele de rezervă din semințe ca alergeni

Unele globuline de rezervă ca ingrediente naturale ale alimentelor pot duce la dezvoltarea sensibilității alergice crescute, având ca rezultat formarea anticorpilor anti-imunoglobulinei E (IgE) specifici alergenilor. Conform datelor prezentate în SDAP (Structural Database of Allergenic Proteins) [3], ambele globuline 7S și 11S sunt alergeni.

Alergenicitatea globulinelor 7S. A fost stabilită alergenicitatea unui număr relativ mare de globuline 7S din semințele utilizate în alimentație (Tabelul 1.6).

Printre globulinele 7S, cea mai mare alergenicitate manifestă globulina 7S din arahide, Ara h1 [79]. Ara h1 este un alergen amjor, care provoacă o sensibilizare mediată de IgE până la 95% dintre pacienți. Acest alergen este o glicoproteină 7S depozitată în semințe care cuprinde 12-16% din conținutul total de proteine din extractele de arahide. Concentrația Ara h1 în arahide crește odată cu mărimea nucleului (4-16 mg extras Ara h1/g arahide), astfel încât expresia proteinei este asociată cu maturitatea arahidelor.

Subunitățile Ara h1 apartin globulinelor 7S de tipul convicilinelor, ale căror secvențe conțin prelungirea variabilă N-terminală în domeniul N-terminal. O parte din această prelungire în molecula matură a Ara h1 este eliminată în timpul procesingului proteolitic post-translațional [13].

În subunitatea globulinei 7S din arahide, Ara h1, a fost stabilită prezența a 21 de epitopi IgE [77]. Epitopii IgE 1-3 aparțin secvenței N-terminale eliminată post-translațional. Dintre cei 18 determinanți antigenici (epitopii IgE), identificați în molecula matură a Ara h1 [77], șase aparțin extensiei N-terminale *a* (Figura. 1.14), sensibilă la proteoliză limitată a globulinelor 7S [70]. Aceasta arată posibilitatea potențială de reducere a nivelului de alergenicitate a Ara h1 prin proteoliză

limitată. Nu este exclus, ca este potențial sensibilă la proteoliză și regiunea extinsă nestructurată *b*, unde este identificat încă un epitop IgE al Ara h1 (Figura 1.14).

Nomenclatura SDAP	Subunitatea	Specia
Ara h1* [77]	P43238	4
Ara h1	P43237	Arachis nypogaea
Gly m α	18536	
Gly m α'	Q9FZP9	Glycine max
Gly m β	256427	
Len c 1.0101* [78]	29539109	Long oulinguig
Len c 1.0102	29539111	Lens cultuaris
Vig r 2.0101	Q198W3	Viewerundinte
Vig r 2.0201	B1NPN8	vigna raaiaia
Lup an 1.0101	169950562	Lupinus angustifolius
Jug r 2	6580762	Juglans regia
Jug n 2	31321844	Juglans niger
Cor a 11	19338639	Corylus avellana
Ses i 3	13183177	Sesamum indicum
Pis v 3.0101	133711973	Pistacia vera
Ana o 1.0101	21914823	An acquidium cooidental.
Ana o 1.0102	21666498	Anacaratum occidentale

Tabelul 1.6. Subunitățile globulinelor 7S care sunt alergeni (conform SDAP).

Notă: \* – subunitățile globulinelor 7S în secvențele cărora au fost identificați determinanții antigen (epitopii IgE)



#### Fig.1.14. Poziția epitopilor IgE în structura terțiară a subunităților globulinei 78 din arahide, Ara h1 [77]

Linia de sus arată structurile secundare (β-strendurile *Z-J'* și α-helixurile *h0-h4*) și regiunile nestructurate (a, b) în secvențele subunităților. Cifrele corespund poziției epitopilor IgE 4-21, identificați în molecula matură Ara h1.

Toți cei patru epitopi IgE, identificați în globulina 7S din linte *Lens culinaris*, aparțin exclusiv regiunii C-terminale a domeniului C-terminal [78].

Alergenicitatea globulinelor 11S. În semințele multor plante de cultură, utilizate în alimentație, globuline 11S sunt alergeni. Potrivit SDAP (Structural Database of Allergenic Proteins), globulinele 11S din 12 plante, inclusiv prune, nuci și hrișcă, sunt alergeni (Tabelul 1.7).

Nomenclatura SDAP	Subunitatea	Specia
Ara h3* [80]	3703107	4
Ara h4 [82]	5712199	Arachis nypogaea
Gly m G1* [14]	169973	
Gly m G2* [81, 83]	295800	
Gly m G3	P11828	Glycine max
Gly m G4	Q7GC77	
Gly m G5	Q9SB11	
Pru du 6 [84]	258588247	
Pru du 6.0101	307159112	Prunus dulcis
Pru du 6.0201	307159114	
Cor a9 [85]	18479082	Corylus avellana
Sin a 2.0101 [86]	Q2TLWQ	Sinapis alba
Car i [87]	158998780	Carya illinoinensis
Jug r 4.0101	Q2TPW5	Juglans regia
Ses i 6.0101	Q9XHP0	Communication and in success
Ses i 7.0101	Q9AUD2	Sesamum inalcum
Ana o2 [88]	25991543	Anacardium occidentale
Fag e1* [89]	2317674	
Fag e1	2317670	Fagopyrum esculentum
Fag e1	29839419	
Pis v 2.0101	110349082	
Pis v 2.0201	110349084	Pistacia vera
Pis v 5.0101	171853009	
Ber e2	30313867	Bertholletia excelsa

Tabelul 1.7. Subunitățile globulinelor 11S, care sunt alergeni(conform datelor prezentate în SDAP).

Nota: \* - indică subunitățile globulinelor 11S, în secvențele cărora sunt identificați epitopii IgE.

Cele mai alergene sunt proteinele din arahide, printre care globulina 11S Ara h3 [80]. În Ara h3, globulina 11S din arahide, au fost identificați patru epitopi IgE [80] (Figura 1.15): primul epitop în regiunea N-terminală a  $\beta$ -barrelui ( $\beta$ -strandul A și  $\alpha$ -helixul h0), al doilea și al treilea epitop, respectiv în regiunea  $\alpha$ -helixului h1/h2 și în regiunea adiacentă  $\beta$ -strendului J' și al patrulea epitop în porțiunea hidrofilă C-terminală.

Printre globuline 11S din cele 12 specii, a căror alergenicitate a fost stabilită, cea mai înaltă alergenicitate după arahide o are glicinina, globulina 11S din soia. Molecula heterohexameră a glicininei din soia este formată din subunități de tipul I (G1-G3) și II (G4, G5), care diferă după structura primară.

În Gly m G1, globulina 11S din soia, au fost identificați doar doi epitopi IgE, a căror poziție se suprapune parțial cu poziția epitopilor 2 și 3 ai globulinei 11S din arahide [14]. Poziția unuia din cei trei epitopi IgE ai Gly m G2 [81], globulina 11S din soia, de asemenea, se suprapune parțial cu poziția epitopului 3 al globulinei 11S din arahide.

Epitopii IgE, identificați în studiile independente ale subunităților globulinelor 11S din arahide, Ara h3 [80] și din soia Gly m G1 [14], și Gly m G2 [81], aparțin  $\alpha$ -catenelor (Figura 1.15). O parte dintre dintre acești epitopi IgE se găsesc în regiunea structurală a  $\alpha$ -helixurilor *h1* și *h2* și  $\beta$ -strendului *J'*. Această regiune în globulinele 11S este sensibilă la proteoliză (Tabelul 1.3), cea ce prezintă posibilitatea principală de reducere parțială a alergenicității globulinelor 11S din arahide și din soia prin proteoliza limitată.

а	b	с	
-Z-A'A-h0-B-C-D-E-E'F'	'F-G-H-I-J-h1'h1-h2·	-h3-J'	
-1-	2	-34-	Ara h3
	1	-2-	Gly m G1
-12-	-34-	-5-	Gly m G2

#### Fig. 1.15. Poziția epitopilor IgE în structurile terțiare ale subunităților globulinelor 11S din arahide, Ara h3 [80] și soia Gly m G1 [14] și Gly m G2 [81]

Linia de sus arată structurile secundare ( $\beta$ -strendurile *Z-J'* și  $\alpha$ -helixurile *h0-h3*) și regiunile nestructurate (a, b, c) în secvențele subunităților. Cifrele corespund poziției epitopilor IgE.

Trebuie remarcat faptul, că o parte din epitopii IgE din Gly m G2 [81] sunt prezenți, de asemenea, în β-catenele globulinelor 11S, insensibile la proteoliza limitată; în același loc, au fost identificați epitopii IgE în subunitățile Fag e1, globulină 11S din hrișcă *Fagopyrum esculentum* [89]. A fost stabilită alergenicitatea tuturor celor cinci subunități ale globulinei 11S din arahide (Tabelul 1.5). Molecula hetero-hexamerică a globulinei 11S din arahide se formează printr-o combinație aleatorie a nouă subunități conservative (Tabelul 1.8); alergenicitatea a două dintre ele este stabilită (Tabelul 1.7). Este foarte probabil ca epitopii IgE identificați în subunitatea Ara h3 (pdb|3c3v) să fie prezenți în regiunile omoloage ale secvențelor, cel puțin, celor mai multe alte subunități.

Subunitatea	Tipul	αβ, kDa	α, kDa	β, kDa
aam46958	Ι	59.696	38.835	20.879
aar02860		59.457	38.837	20.638
aad47382	II	58.849	37.929	20.938
abg93402		58.574	37.921	20.671
3c3v		58.485	37.963	20.540
aag01363		58.223	37.638	20.603
abl14270		58.214	37.570	20.662
adq53859	III	56.284	35.614	20.688
aau21492	IV	52.513	31.973	20.558

Tabelul 1.8. Subunitățile de tipul I-IV, ce diferă după masa moleculară (kDa), din molecula heterohexamerică a globulinei 11S din arahide (conform datelor prezentate în NCBI).

Subunitățile atât ale globulinelor 7S, cât și ale globulinelor 11S sunt foarte conservative, ceea ce determină legitățile comune ale proteolizei lor limitate. Probabil, proteoliza limitată poate fi utilizată pentru a reduce alergenicitatea nu numai a câtorva subunități ale globulinelor din arahide și din soia, studiate mai sus, ci și a altor subunități ale acestor proteine, precum și a globulinelor de rezervă din semințele altor plante, alergenicitatea cărora a fost stabilită. Programul SDAP oferă capacitatea de a evalua cantitativ probabilitatea prezenței epitopilor IgE, identificați în globulinele 7S și 11S din arahide și din soia [14, 77, 80, 81], în regiunile omoloage ale secvențelor globulinelor de rezervă din semințele altor plante. Cu toate acestea, astfel de cercetare nu a fost efectuată.

#### 1.5. Concluzii la capitolul 1.

Analiza tematică la subiectul tezei ne permite sa facem unele concluzii:

1. Precursorii evolutivi străvechi ai globulinelor de rezervă din semințe pot fi fie oxalatdecarboxilazele bacteriene OD, fie sferulinele Amoebozoa, înrudite cu globulinele de rezervă după structurile primare, terțiare și cuaternare. Până în prezent, nu există un răspuns convingător la această întrebare.

2. Globulinele de rezervă 7S (vicilinele) și globulinele 11S (leguminele) sunt cele mai tipice pentru semințe proteine de rezervă, ale căror precursori evolutivi direcți sunt proteinele vicilino- și legumino-similare din plantelor cu spori. Este evident că structurile primare și superioare atât ale globulinelor 7S, cât și ale globulinelor 11S reflectă cele mai vechi trăsături ale predecesorului lor evolutiv comun. Această problemă este încă deschisă.

3. S-a constatat că degradarea masivă a unor globuline de rezervă din semințe *in vitro* și din semințele germinate, care decurge ca o reacție de ordinul pseudo-unu, începe doar după finalizarea proteolizei limitate site-specifice anterioare. Evident, în timpul proteolizei limitate, apar modificări structurale ale globulinelor de rezervă, expunând legăturile peptidice mascate în structura lor nativă. Natura acestor schimbări, care redau globulinelor de rezervă abilitatea de a degrada profund, nu este descrisă.

4. Rolul proteolizei limitate în inițierea degradării masive a globulinelor 11S este stabilit, în mod credibil, pentru globulinele din semințele a numai două plante (soia şi dovleac). Rămâne neclar în ce măsură acest rol este caracteristic întregii familii de globuline 11S.

5. Rolul proteolizei limitate în inițierea degradării masive a globulinelor 7S este stabilit numai pentru fazeolină. Nu este clar dacă un astfel de rol este caracteristic și pentru alte proteine din familia globulinelor 7S.

6. S-a stabilit alergenicitatea multor globuline de rezervă 7S și 11S din semințe. În structura unora dintre ele, au fost identificați determinanții antigenici (epitopi IgE). O parte dintre acești epitopi IgE aparțin regiunilor structurilor globulinelor 7S și 11S, potențial sensibile la proteoliza limitată. Aceasta arată principala posibilitate de reducere parțială a nivelului de alergenicitate a globulinelor de rezervă din semințe prin proteoliza limitată. Cu toate acestea, cercetări în această direcție nu au fost efectuate.

## 2. ORIGINEA ȘI EVOLUȚIA STRUCTURILOR GLOBULINELOR DE REZERVĂ DIN SEMINȚE

Conform rezultatelor analizei evolutive întreprinse de cele două echipe de cercetători de la universitățile din Marea Britanie [1,5,46-49] și din Moldova [2, 4, 6], predecesorii evolutivi străvechi ai globulinelor de rezervă din semințe pot fi fie OD bacteriene, fie germinele vegetale și precursorii necondiționați ai acestora, sferulinele din Amoebozoa (capitolul 1). Astfel, problema originii globulinelor de rezervă rămâne deschisă. Această problemă poate fi soluționată utilizând o strategie adecvată de cercetare a relațiilor evolutive între cele trei grupuri de proteine, pentru care este stabilită relația dintre structurile primare și cele superioare – globulinele de rezervă, OD și germinele. Principalele elemente ale unei astfel de strategii, descrise mai sus [6, 89], sunt completate și dezvoltate în această lucrare.

#### 2.1. Metode și strategia de cercetare

În lucrare au fost utilizate următoarele programe:

• BLAST (<u>http://www.ncbi.nih.gov/</u>) – pentru căutarea secvențelor omoloage de aminoacizi ale proteinelor;

- Clustal W2 și ulterior Clustal Omega pentru alinierea (alignment) secvențelor lor;
- TREECON [90] pentru analiza lor evolutivă;

• DeepView/Viewer Swiss-PDB – pentru analiza structurilor terțiare ale proteinelor, construirea diagramelor panglică și alinierea structurală. Asemănarea structurilor terțiare și cuaternare a fost evaluată prin rezultatele suprapunerii spațiale a acestora. Au fost acceptate doar opțiunile oferite de program în care atomii  $C^{\alpha}$  ai resturilor de aminoacizi sunt suprapuși cu o abatere standard de cel mult 1,6 Å. Numărul resturilor de aminoacizi identici în secvențele subunităților aliniate structural și numărul atomilor  $C^{\alpha}$  suprapuși spațial în subunitățile individuale și dimerii acestora au fost luate ca criterii pentru similaritatea structurilor primare, terțiare și cuaternare a proteinelor oligomere.

• Swiss Model [91, 92] – pentru a construi structurile terțiare și cuaternare model ale proteinelor studiate.

Pentru a obține din punct de vedere statistic o secvență de încredere de formare a clusterelor pe arborele evolutiv al globulinelor de rezervă, OD și germinelor, am fost ghidați de următoarea strategie.

1. Crearea colecțiilor independente de proteine, care formează clusterele globulinelor de rezervă 11S și 7S, OD și germinelor și predecesorilor fie necondiționați sau cei mai probabili ale fiecăreia dintre aceste patru colecții. Utilizarea structurile primare, care corespund celor patru predecesori, ca secvență-rădăcină (outgroup sequences) pentru a construi patru arbori evolutivi independenți. Din setul preliminar vast de sute și chiar mii de secvențe, ale fiecărei din cele patru grupe trebuie să fie alese numai acelea, ce ocupă poziția cea mai apropiată de trunchiul principal al fiecăruia dintre cei patru arbori evolutivi; acest lucru ar trebui să fie indicat de valorile suficient de ridicate ale suportului statistic al clusterelor (bootstrap values). Pe de altă parte, în cadrul fiecăruia dintre cele patru grupuri trebuie să fie prezentată toată diversitatea taxonomică a speciilor, la care aparțin proteinelor ce urmează să fie analizate.

2. Segmentul de secvențe analizat ar trebui, pe cât posibil, să fie mai lung. Alinierea secvențelor din setul preliminar vast de proteine ale fiecărei din cele patru grupe (inclusiv, cu structura terțiară cunoscută) arată, că această regiune cuprinde, cel puțin, întreaga secvență a modulului structural  $\beta$ -barrel –  $\alpha$ -helixuri (inclusiv domeniile N- și C- terminale ale proteinelor bidomenice).

3. Conform datelor din literatură [2, 38, 41, 43] și, de asemenea, în conformitate cu rezultatele prezentei lucrări, domeniile C-terminale nu numai ale globulinelor 11S, dar și ale globulinelor 7S sunt cele mai conservative și în cea mai mare măsură reflectă cele mai vechi trăsături ale proteinelor de rezervă. Prin urmare, este oportun ca pentru analizele evolutive să se utilizeze numai secvențele domeniilor C-terminale ale globulinelor de rezervă și, respectiv, domeniile C-terminale omoloage ale proteinelor bidomenice cu alte funcții (non-rezervă).

4. În cazul alinierii secvențelor proteinelor, selectate în conformitate cu alineatul (1), în unele dintre acestea sunt detectate inserții scurte individuale. Astfel de inserții nu influențează la formarea arborelui evolutiv în cadrul uneia din cele patru grupuri, fiecare dintre ele este caracterizat printr-un nivel ridicat de conservatism. Cu toate acestea, la fuzionarea grupurilor (etapa 5) prezența acestor inserții duce la formarea unor porțiuni inadecvate ale alinierii, care sunt depistate a) prin nesuprapunerea poziției structurilor secundare cunoscute sau prezise ale proteinelor grupurilor comparate și b) prin valori joase ale suportului statistic al clusterelor. Prin

urmare inserțiile, specifice pentru o singură sau un număr limitat de secvențe dintr-un grup, trebuie să fie eliminate înainte de unirea grupurilor. Aceste inserții pot fi readuse în poziția inițială a alinierii celor patru grupuri unite, după care poziția lor poate fi ajustată prin analizarea zonei limitate corespunzătoare alinierii finale.

5. Unirea grupurilor de globuline de rezervă 11S și 7S și construirea arborelui evolutiv corespunzător. Formarea unui cluster unificat, care conține toate cele trei clustere inițiale – globulinele de rezervă, OD și germinele, și construirea arborelui fără rădăcină corespunzător.

6. În acest arbore fără rădăcină, una dintre secvențe se dovedește a fi cea mai puțin similară cu secvențele din clusterele OD și germinelor. Anume această secvență, ce poartă trăsăturile predecesorului evolutiv atât al OD, cât și al germinelor, ar trebui să fie folosită în calitate de rădăcină pentru construirea arborelui evolutiv care finalizează studiul.

#### 2.2. Analiza originii și evoluției globulinelor de rezervă

**Căile independente ale evoluției globulinelor de rezervă și germinelor.** Strategia descrisă mai sus, a fost utilizată pentru a identifica predecesorii străvechi evolutivi ai globulinelor de rezervă din semințe și pentru a descrie etapele succesive ale evoluției lor. Folosind programul BLAST, (Basic Local Alignment Search Tool, (instrument de bază de căutare a alinierii locale) <u>http://www.ncbi.nih.gov/</u>), s-a format o vastă colecție de secvențe de aminoacizi omoloage ale proteinelor din suprafamilia globulinelor de rezervă din semințe: globulinelor de rezervă propriuzise, proteinelor din plantele cu spori și din algele verzi, oxalatdecarboxilazelor bacteriene, precum și germinelor și sferulinelor și proteinelor înrudite cu ele din plante și din eucariotele inferioare.

A fost efectuată alinierea multiplă a secvențelor de aminoacizi ale proteinelor suprafamiliei globulinelor de rezervă (programele Clustal W2 și Clustal Omega) și pe baza ei s-a construit arborele fără rădădcină (programul TREECON) care descrie relațiile filogenetice dintre proteinele studiate.

În arborele fără rădăcină (Figura 2.1), care unește colecția selectată de proteine studiate (Tabelul 2.1), proteina monodomenică cam64834, calificată ca oxalatdecarboxilaza ipotetică din actinobacteria *Mycobacterium abscessus*, arată o similitudine cu domeniile C-terminale ale OD bacteriene (Figura 2.2). În acest caz, însă secvența completă cam64834, cu mult peste modulul structural  $\beta$ -barrel- $\alpha$ -helixuri, este omoloagă sferulinelor monodomenice din Amoebozoa (Figura 2.2). Fără îndoială, secvența acestei proteine, care reflectă trăsăturile vechi atât ale oxalat-decarboxilazelor, cât și ale sferulinelor [6] poate fi folosită în calitate de rădăcină la construirea arborelui, care descrie originea și evoluția globulinelor de rezervă din semințe (Figura 2.3).



Fig. 2.1. Arborele fără rădăcină, care descrie relația filogenetică dintre clusterele globulinelor de rezervă, OD și germinelor

Numerele de deasupra ramurilor reflectă suportul statistic al clusterelor (% din 1000 de replicări). Săgeata indică proteina OD-similară monodomenică cam64834 din *Mycobacterium abscessus*, care ocupă o poziție intermediară între clusterle OD bacteriene bidomenice și germine/sferuline.

Codul	Specia	Taxonul	Proteina
aak15089	Sesamum indicum	Eudicotyledons	
ern01581	Amborella trichopoda	basal Magnoliophyta	Visiling
caa44873	Picea glauca	Coniferopsida	Viciline
caa90652	Zamia furfuracea	Cycadophyta	
baa06186	Cucurbita maxima	Eudicotyledons	
ern04969	Amborella trichopoda	basal Magnoliophyta	
efj36753	Selaginella moellendorffii	Lycopodiophyta	Vicilino-similare
efj09997	Selaginella moellendorffii	Lycopodiophyta	
caa91187	Matteuccia struthiopteris	Moniliformopses	
caa57848	Magnolia salicifolia	Magnoliids	
ern18747	Amborella trichopoda	basal Magnoliophyta	T
caa44874	Picea glauca	Coniferopsida	Legumine
caa90641	Ginkgo biloba	Ginkgophyta	
efj09931	Selaginella moellendorffii	Lycopodiophyta	Legumino-
efj10701	Selaginella moellendorffii	Lycopodiophyta	similare <i>a</i>
ern17205	Amborella trichopoda	basal Magnoliophyta	
ade77789	Picea glauca	Coniferopsida	
efj36160	Selaginella moellendorffii	Lycopodiophyta	Legumino-
efj18490	Selaginella moellendorffii	Lycopodiophyta	
efj19431	Selaginella moellendorffii	Lycopodiophyta	
eie22888	Coccomyxa subellipsoidea	Chlorophyta	OD similars
eie22877	Coccomyxa subellipsoidea	Chlorophyta	- OD-similare
acb94629	Beijerinckia indica	Proteobacteria	
ack49092	Methylocella silvestris	Proteobacteria	
aie72611	Synechocystis sp.	Cyanobacteria	
2vqa	Synechocystis sp.	Cyanobacteria	
aac04835	Oryza sativa	Liliopsida	Carrier
ern02090	Amborella trichopoda	basal Magnoliophyta	Germine
abk22930	Picea glauca	Coniferopsida	
efj08377	Selaginella moellendorffii	Lycopodiophyta	
bad86501	Physcomitrella patens	Bryophyta	Germino-similare
eie27072	Coccomyxa subellipsoideas	Chlorophyta	1
cbn75542	Ectocarpus siliculosus	Stramenopiles	
aaa29979	Physarum polycephalum	Amoebozoa	Steruline
cam64834	Mycobacterium abscessus	Actinobacteria	OD-similare

# Tabelul 2.1. Taxonomia organismelor, ale căror proteine sunt folosite pentru a analizarelațiile evolutive dintre globulinele de rezervă și germine

1 MOVRNILVALVVVCFAVSEAA-TQAPTSPPDSSASAEQVAQLLNAPS PHYpo 4 INRRTVLGGVGVAGVAAAAGAGTDWAFSPPRSHDNPEDVTVTDDAAR MYCab ·· \* ·· \* · · \* · · \* \* · · \* \* \* \* · · \* \* \* . . \* : \* : • \* . А РНҮро 47 ELD-RLKLLKDNQFVFDFKNSKLGVTQSAGGKTVATSRTDFPAVIGH 52 FGDPRLPAELNTAQPHLFHLGALAPQTFDGGDLRQAHEGNFPILTGQ MYCab \* \*\* :. . \*: . \*. \*\*. :. :\*\* : \*: C' С D F В E PHYpo 93 NIAMTVGFIEACGINLPHTHPRATEINFIAKGRFQAGFFLENQATFI 99 QASVVMVTLQPGGIREPHWHPSAWEINVITSG-VATWTLLDPMGHSE MYCab G Η Т J PHYpo 140 GHILEEGMATVFPQGAIHFEINLECEPAMFVAAFNN---EDPGVQTT MYCab 145 TFDAHVGDVVFAPQGSLHYFENKGPDDLKLLIVFNASTAEGKDDIGI • \* ••• \*\*\*• \* :: .\*\* \*... : h1 h3 PHYpo 184 ASSFFGLPADVAAVSLNISSIQTVEDLAKYLPHNPAIAMKECMQR MYCab 192 GASISKLPPDVLSAVFGVP----TETFAKFKKINESVTILRRPOT .\* :\*\*: \* :::: \*

#### Fig. 2.2. Alinierea secvențelor de aminoacizi ale sferulinei aaa29979 din Amoebozoa Physarum polycephalum (PHYpo) și ale proteinei OD-similare cam64834 din actinobacteria Mycobacterium abcesus (MYCab)

Resturile de aminoacizi identice sunt evidențiate cu bold. Structurile secundare ale sferulinei au fost obținute în lucrarea de față prin modelarea omoloagă a structurii sale terțiare (în calitate de sablon germina pdb|1fi2 din *Hordeum vulgare*).

Așa cum s-a presupus anterior [6], globulinele de rezervă și germinele formează două ramuri evolutive independente, provenind din proteina OD-similară monodomenică cam64834 din actinobacteria *Mycobacterium abcessus* (Figura 2.3). Evoluția germinelor s-a dezvoltat după scenariul divergent clasic (Bacteria, Amoebozoa, Stramenopiles, Chlorophyta, Bryophyta, Lycopodiophyta, Pinophyta, Magnoliophyta).

Dimpotrivă, pe ramura evolutivă a globulinelor de rezervă, lipsesc proteinele bidomenice din eucariotele inferioare ca intermediari între proteinele OD bacteriene și proteinele OD-similare din alga verde *Coccomyxa subellipsoidea*. Sunt posibile două căi ale evoluției globulinelor de rezervă, care explică acest fenomen.

*Ipoteza 1.* S-a demonstrat, că multe mii de gene au fost transferate din genomul endosimbiontului cianobacterial în genomul nuclear al plantelor [93-95]. Existența genelor algei verzi *Coccomyxa subellipsoidea*, care codifică proteinele OD-similare poate fi considerată ca o ilustrare selectată cu succes a implicării endosimbiozelor în evoluția globulinelor de rezervă din semințe.



# Fig. 2.3. Căile independente ale evoluției globulinelor de rezervă din semințe și a germinelor [96]

Regiunile secvențelor, utilizate pentru analiză (Tabelul 2.1), corespund domeniilor C-terminale ale proteinelor bidomenice și domeniului germinelor/sferulinelor (172 de poziții ale alinierii multiple, care cuprind modulul structural β-barrel – α-helixuri). Numerele de deasupra ramurilor reflectă suportul statistic al clusterelor (% din 1000 de replicări). Linia punctată separă clusterele proteinelor monodomenice și bidomecice. Ca rădăcină, a fost utilizată secvența proteinei ODsimilare monodomenice cam64834 din *Mycobacterium abscessus*. *Ipoteza 2.* Proteinele OD-similare bidomenice din alga verde *Coccomyxa subellipsoidea* sunt direct moștenite de la OD bacteriene bidomenice [46], posibil, cu participarea mecanismului transferului orizontal de gene. În acest context, evidențiem relația apropiată care am descoperit-o între proteinele OD-similare bidomenice din Metazoa și OD bacteriene, în absența unor legături evolutive intermediare cunoscute între ele (datele nu sunt prezentate).

Analiza relațiilor filogenetice între OD cianobacteriene și un grup de OD proteobacteriene înrudite, care formează un cluster comun (Figura 2.3), nu a permis de a alege între ipotezele alternative descrise [6].

Cu toate acestea, următoarele observații susțin ipoteza 2. OD a fost găsită printre proteinele a numai două genuri de cianobacterii (*Synechocystis* și *Synechococcus*) și numai în cinci (RCC6301, RCC6312, RCC6714, RCC6803 și RCC7942) din 54 de genomuri de cianobacterii [97]. Astfel, genele care codifică OD nu sunt caracteristice pentru cianobacterii. Poate că genele proteinelor OD-similare au apărut într-un număr limitat de genomuri de cianobacterii relativ recent, cu participarea mecanismului transferului orizontal de gene.

Următoarea informație detaliază calea evolutivă a globulinelor de rezervă din semințe. În genomul *Selaginella moellendorffii* (struțișor) și *Amborella trichopoda* (grupul bazal Magnoliophyta), există gene, care codifică toate cele cinci tipuri de proteine omoloage cu vicilina și legumina: proteinele vicilinosimilare, proteinele leguminosimilare din grupul *a* și *b*, germinele și proteinele germinosimilare (Figura 2.3).

Proteinele din *Selaginella moellendorffii* (vicilino-similare și legumino-similare din grupul *b*) sunt precursorii nu doar ai globulinelor de rezervă din semințe, dar, și a două mari familii conservative de proteine vicilino-similare și legumino-similare din spermatofite, cu alte funcții (funcții non-rezervă) (Figura 2.3). Cu excepții rare [4, 43], funcțiile acestor proteine rămân necunoscute.

Proteina vicilino-similară caa91187 este sintetizată în mod special în sporii de maturare ai ferigii *Matteuccia struthiopteris* [98]. S-a arătat că mecanismele de reglare a expresiei genei sale și a genelor globulinelor de rezervă din semințe sunt similare [99]. Probabil, proteina vicilinosimilară din feriga *Matteuccia struthiopteris* este o proteină de rezervă primitivă [100]. Poate că, proteinele vicilino-similare și legumino-similare din grupul *a* din struțișor *Selaginella moellendorffii*, care formează un cluster comun cu leguminele, sunt proteine de rezervă primitive (Figura 2.3).

Subunitățile globulinelor de rezervă și OD sunt combinate în trimeri, formați în rezultatul interacțiunilor dintre domeniile N- și C-terminale ale subunităților adiacente în trimeri. Pe parcursul evoluției leguminelor s-a format un site-procesing Asn-specific interdomenic, al cărui

clivaj post-translațional duce la formarea  $\alpha$ - și  $\beta$ -catenelor ale subunităților (respectiv, domeniile N- și C-terminale).

**Proteinele legumino-similare din alga verde** *Klebsormidium nitens*. Până de curând, era neclar, daca genele, care codifică nu numai proteinele OD-similare din alga verde *Coccomyxa subellipsoidea*, dar și proteinele legumino- și vicilinosimilare, sunt caracteristice genomurilor altor alge verzi. Abia la sfârșitul lunii mai 2017 a fost raportată informația despre genomul algei verzi *Klebsormidium nitens*. Printre proteinele ipotetice ale acestei alge BLAST căutarea dezvăluie cel puțin patru secvențe de aminoacizi conservative ale proteinelor bidomenice (gaq91246, gaq85384, gaq84499 și gaq84498), omoloage leguminelor din semințe (Figura 2.4) și, prin urmare, proteinelor leguminosimilare din struțișor *Selaginella moellendorffii*.

			Α'	А	h0	В	С	D
gaq91246	16	AGVCCGMD-LVPRGPTE	ILKTEGG	CIKQWEP	DSALGHYAA	AGAVEVTL	EPQGVLLPRYN	MTPSLLYV
3qac	10	qGNECQIDRLTALEPTN	RIQAERG	LTEVWDSI	NEQEFRCAG	VSVIRRTI	EPHGLLLPSF	<b>FSAPELIYI</b>
		* * *. **:	:::* *	: *:	:. : *.	:. *:	**:*:***	.:*.*:*:
		F					F	G
aaa91246	82							
3020 3020	77		Vasasaa	faaada	rirpaasrk	famradrf	UDOHOKIBHI	RECOTEAMP
Sque	, ,	<u>.*</u> *	resysyy	rqggcac.	a a	.rymryarr	••••*	••**• *•*
		• • • • • •			u		••••	•••••
		Н	I		J			
gaq91246	112	EGF <u>VLGVHN</u> DGQDK <u>FRA</u>	LGVADLS	NTRGKGAI	MN-A <u>SIFFI</u>	<u>F</u> GN		-REDETLAP
3qac	145	AGV <u>SHWAYN</u> NGDQP <u>LVA</u>	VILIDTA	NHANQLDI	KNFPTR <u>FYI</u>	AGKPQQEH	Sgehqfsresi	rrgerNTGN
		*:*:*:: : *	::*:	* .:	* : *:*	*:	b	* :
		h1 h2	h3		J'			
gag91246	162	VINGFSKDVLAAAWNVD	EATVEKL	LTSQKGA	GIIK-VEKK	ESARAPA-		
3gac	211	IFRGFETRLLAESFGVS	EEIAOKL	OAEODDR	GNIVRVOEq	lhvikpps	raweereqqsi	rgsrylpn↓
-		:**	* .:**	<u>.</u> *	* * *::	. *:	c	5 1 1
		2	7	/ 7		10	D	0
	200						B	
gaq91246	208	EPEPNTFFGD <u>FSFNL</u> EH	TQAEV	VIRGGSL	<u>NEV</u> TKEKLE	TTELLOTO	AAHTKLKPDAI	LEAPTWSAN
Sqac	211	gveeticsar <u>lavnv</u> dd	PSKAD <u>VI</u>	TPEAGRE	<u>+</u> +.++	<u>11441</u> 812		MAPHINLN
		* ::.*::.		••••	• ••••	· · · · · ·	**: * :*:	
		D E		F	G	Н	I	J
gaq91246	273	AHHIVYFTRGKGRIQVA	<u>g</u> pgg <u>ktv</u>	LDRAVGE	G <u>DLVVI</u> PKH	FPAAKLAA	RAAPLEWVDVN	MTSPAPLSF
3qac	304	AHNIMYCVRGRGRIQIV	NDQGQSV	FDEELSR	GQLVVVPQN	FAIVKQAF	-EDGFEWVSFH	<u>KT</u> SENAMFQ
		**:*:* .**:***:.	. *::*	:*. :	*:***:*::	* .* *	:***	** :
		h1 h	2	h3		J <b>'</b>		
gag91246	340	FLAGANSVYKAFPFDVL	VAGFNVD	PEVEKKV	LHARAHEAM	IILTRP		
3gac	410	SLAGRTSAIRSLPIDVV	SNIYOIS	REEAFGL		FRSSG		
T		*** *	:::.	* :	_ * . <u>::</u>	:		

#### Fig. 2.4. Alinierea secvențelor de aminoacizi ale proteinei gaq91246 din *Klebsormidium nitens* și globulinei 11S pdb|3qac din *Amaranthus hypochondriacus*

Aici şi în continuare structura pdb|3qac este utilizată ca matriță pentru modelarea structurii terțiare şi cuaternare a proteinei gaq91246. Poziția structurilor secundare (β-strendurile *Z-J'* şi α-helixurile *h0-h3*) este indicată prin subliniere. Inserțiile extinse *a*, *b* şi *c*, specifice secvențelor globulinelor 11S [102], în proteina legumino-similară din alga verde *Klebsormidium nitens* sunt absente. Săgeata indică hotarul dintre α- şi β-catenele (domeniile N- şi C-terminale) globulinei 11S.

În structura terțiară model a subunității proteinei gaq91246 din alga verde *Klebsormidium nitens* există două domenii care formează aproape toate structurile secundare caracteristice subunităților leguminelor (Figura 2.4 și 2.5). La fel, ca și în cazul proteinelor legumino-similare din struțișor *Selaginella moellendorffii*, în secvența domeniului N-terminal din alga verde *Klebsormidium nitens* lipsesc inserțiile extinse *a* (între  $\beta$ -strendurile *E* și *F*) și *b* (între  $\beta$ -barrel și  $\alpha$ -helixuri); linkerul interdomenic *c* relativ scurt, așa cum este tipic pentru proteina OD-similară din alga verde *Coccomyxa subellipsoidea* și OD bacteriană 2vqa (Figura 1.10).



Fig. 2.5. Diagramele panglică ale structurilor terțiare ale domenilor C-terminale ale proteinei gaq91246 din alga verde *Klebsormidium nitens* (a, model) și globulinei 118 pdb|3qac din amarant *Amaranthus hypochondriacus* (b), suprapuse spațial, dar prezentate separat.

În structura cuaternară model a proteinei gaq91246 din alga verde *Klebsormidium nitens*, subunitățile sale (*a*, *b*, *c*), în același mod ca și în globulinele 11S din semințe și în OD bacteriene [96], sunt combinate în trimerii *abc* prin interacțiunile între regiunile C-terminale ale domeniilor

N-terminale și C-terminale ale fiecăreia dintre cele trei perechi de subunități adiacente în trimeri (Figura 2.6).



Fig. 2.6. Structura cuaternară model a proteinei legumino-similară gaq91246 din alga verde *Klebsormidium nitens a*, b, c – subunitățile proteinei; N și C – domeniul N- și C- terminal, respectiv.

Într-o BLAST căutare pe scară largă, folosind în calitate de șabloane secvențele nu doar a globulinelor 7S din semințe, dar și a proteinelor vicilino-similare din plantele cu spori, nu s-au depistat proteinele vicilino-similare în alga verde *Klebsormidium nitens*. Aceasta indică antichitatea relativă a structurilor primare ale globulinelor 11S comparativ cu globulinele 7S. Aceeași concluzie a fost trasată și de rezultatele suprapuneri structurii terțiare a OD bacteriene pdb|2vqa și a tuturor structurilor cristaline cunoscute ale globulinelor 7S și 11S [96].

Rezultatele de mai sus ale suprapunerii structurilor primare și superioare ale globulinei de reazervă 11S pdb|3qac din amarant *Amaranthus hypochondriacus* și proteinei gaq91246 din alga verde *Klebsormidium nitens* (Figura 2.5 și 2.6) permit să presupunem, că proteinele leguminosimilare ale acestei alge verzi reflectă trăsăturile specifice predecesorului străvechi al globulinelor 11S. Pentru a verifica această ipoteză, colecția de secvențe conservative domeniilor C-terminale ale globulinelor de rezervă din semințe, ale proteinelor legumino- și vicilino-similare din plantele cu spori, ale proteinelor OD-similare din algele verzi *Coccomyxa subellipsoidea* și OD bacteriene [96] a fost completată cu secvențele proteinelor legumino-similare gaq91246 și gaq85384 din algele verzi *Klebsormidium nitens* [101]. La alinierea multiplă a secvențelor de aminoacizi ale domeniilor C-terminale ale proteinelor din această colecție s-au depistat numeroase poziții conservative, care conțin resturi de aminoacizi identice (Figura 2.7). Pe baza alinierii multiple a colecției îmbogățite, a fost analizată calea evolutivă a globulinelor de rezervă 7S și 11S, utilizând (ca mai înainte [96]) ca rădăcină secvența OD monodomenică din actinobacteria *Mycobacterium abscessus* (Figura 2.8).

```
aak15089 INIYQQRPT-HSNQYGQLHEVDASQYRQLRDLDLTVSLANITQGAMTAPHYNSKATKIALVVDGEGY
ern01581 FNLFSRRPR-ISNDHGELYELDETEYSPLREFDIAISFANISRGSLEVPFYNSRATEIYVILEGRAR
caa90652 FNLLYRNPD-FSNNNGEIFTADAADHRVLRRLNVGVQLLNLKPRSMTAPHYDTRSTRIGIVRNGRGI
caa44873 FNLRNQKPD-FENENGRFTIAGPKNYPFLDALDVSVGLADLNPGSMTAPSLNSKSTSIGIVTNGEGR
caa91187 YNLFKEKAD-FGNDYGSTTTIHGKDFKLLKALNKGVFLVRLKAGAVLAPHWNPRATEIALVTKGEGE
efj36753 FNLFKKKPD-FENDNGRAITVDGRQYAPLRNASVGVFGVSLKPAAILAPHWNPRAAEIALVTKGQGV
efj09997 FNLLKKAPD-FKNENGWTIAVHGAEFSPLKEADVGVFAVSLKPGAMLAPHWNPRAAEVAFITKGNGR
caa57848 YMDNPREADIYSRQAGRLNSVNMNKLPILRMLGMSSEKGYLYQNAIFSPHWTINAHNIFYVTRGEAR
ern18747 NIDNPRRADIYSROAGHIOIVNROTLPILSILDMSIEKGHLHPNALYAPHWTINAHTIVLITRGEGN
caa90641 NADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILOPNAMFAPSW-LNAHAVMYVTRGOGR
caa44874 NADNPEDADVYVRDGGRLNRVNRFKLPVLKYLRLGAERVVLHPRASCVPSWRMNAHGIMYVTRGEGR
efj10701 YRLGDSQPDVYVPRGGELRQLNSMKMPILKELGLSAACYQLKSGALTAPAWAHNAHKVIYVNEGRGR
efj09931 YRWSHLQPDVRVRDAGELRLLNSFKLPILKKLNMGAAYLKMEAGALTAPGWIQNAHKVMYVERGDGR
gag85384 FNLEETPPD-IVNKGGSYKAVTVSRLPSLEKVDISAAYATLKPDAIFAPSRYFPAHRVTYFTRGKGR
gag91246 FNLEHTQAE-VVTRGGSLNFVTKFKLPILETIGLSAAHTKLKPDALFAPTWSANAHHIVYFTRGKGR
eie22877 FQVSESIPL-IDAAGGWVREASTELFPIASDMSG--AVIRFAPGGLRQIHWHPSLSEWQFVINGTLR
eie22888 FQLAQVIPE-VNTAGGFISIAQVTNFPVSTEMSG--ALVRFAPGALRQLHWHPGHAEWQYVINGSLT
ack49092 FPLGDOTPV-FOSAGGAVRIADSENFPA-TTMAA--ALVEIAPGGMRELHWHPLSDEWOYYISGTAQ
acb94629 FALMDOTPI-KT-RSGTVRIADVNNFKASTNMSA--ALVEVEPGGMRELHWHPNSDEWOYYISGSGR
aie72611 HNLLGOOPL-VSVGGNELRLASAKEFPGSFNMAG--GLIHLEPGAMRRLHWHPNADEWOYVLDGEMD
2vga
         HNLLGQQPL-VSLGGNELRLASAKEFPGSFNMTG--ALIHLEPGAMRQLHWHPNADEWQYVLDGEMD
cam64834 FHLGALAPQ--TFDGGDLRQAHEGNFPILTGQQASVVMVTLQPGGIREPHWHPSAWEINVITSGVAT
          \boldsymbol{z}
                    A/
                         A
                                            B
                                                        C
                                                                     D
```

#### Fig. 2.7. Alinierea secvențelor de aminoacizi, utilizată pentru construirea arborelui evolutiv al superfamiliei de proteine de rezervă (partea inițială)

Linia de jos indică poziția structurilor secundare. Resturile de aminoacizi, care sunt identice cel puțin în jumătate din structurile primare date ale globulinelor de rezervă și ale precursorilor lor, sunt evidențiate cu bold.

cam64834	4 WTLLDPM	MGHS-ETFDAHV	GDVVFAPQGSLH	YFENK <b>G</b> PDD <b>L</b> KLLI	VFNASTAEGK-DDIG	IGASI	ISKL
	E	F	G	H	I	J	h1
aak15089 ern01581 caa90652 caa44873 caa91187	ER <b>E</b> AKELA ERAAKEVS PREIKRLV EREIRQLS EAEVLASS	AFGMPAREVEEVS SFNVPSREAEEVL VFPSSAEEIEATL SFNVPGEEIEEVL SFNIPVEHLQHFL	RSQQEEFFFKG RAQSDSVFVVG EAQEDEVLLNA QAQKDQVILRQ HLQPEAVILPG	Sesamum Amborella Zamia Picea Matteucia	Viciline		
efj36753 efj09997	DE <b>E</b> TLSTA SKDIMSKI	A <b>F</b> AAPPEKLRDFL F <b>F</b> NVDAKRLEAVL	.GWQKDAVILAG .DNQRDAVILPG	Selaginella Selfginella	Vicilinosimilare (stuțișor)		
caa57848 ern18747 caa90641 caa44874	PLEVLINA PIEVLKNA PQEVVMNA PEEVLSAA	AYQVSYREAQNLK AYKITTQEARDIK AYRINGKDARDLR AYRMDRTEVRQIM	FNREHQLMFFP LNRKDQYMLLP RNREHETIILS IRNRERDTLILP	Magnolia Amborella Ginkgo Picea	Legumine		
efj10701 efj09931	PRSVMSSA PRQVLSSA	ALQIDEQLQQKLE A <b>F</b> QIDEKTQQQLE	DRRAEEAYIFP DARSEDAYIFP	Selaginella Selaginella	Leguminosimilare (struțișor)	)	
gaq85384 gaq91246	PFDWLVTC PFD <b>VL</b> VAC	3 <b>F</b> NADPEIEKVVL 3 <b>F</b> NVDPEVEKKVL	RARTDDAFIST HARAHEAMILT	Klebsormidium Klebsormidium	Leguminosimilare (algele verzi)	)	
eie22877 eie22888	PYQVSATT DP <b>E</b> ITAAS	rlnstlafaqs slntsvafaka	-ISYSQPSMAG -INYAVPTMVP	Соссотуха Соссотуха	OD-similare (algele verzi)		
ack49092 acb94629 aie72611 2vqa	PHELVSAH PHELVAAH PNSVLGNT PSSVLGNT	HLNIDVDTLRK HLNMDVSQLRQ FFQISPQLTQQ FFQISPELTKK	-MAQSKTPVVP -LAQNKTPVVP -LPVEDVIFSP -LPVQDTIFSL	Methylocella Beijerinckia Synechococcus Synechcystis	OD (bacterii)		
cam64834	PPDVLSAV h2	V <b>F</b> GVPTETFAKFK <b>h3</b>	KINESVTILRR J'	Mycobacterium	OD (actinobacter	ii)	

LTVFASEGKA-SVSRLQQGDVGYVPKGYGHAIRNSSQKPLDIVVVFNDGDY----QSIDLSTWLASN

2vga

aak15089 FEMACPHSYQRVASRLTRGTVVIIPAGHPFVAVASSNQNLQVLCFEVNANNNEKFPLAGRRNVMNQL ern01581 IEMACPHSYHKIRSDLNPEDLFVAPPSHPIAIRASQEENLQIICFEINARRNRKYFLAGKNSILNQI caa90652 LELVRPQQYQKLRANLNPGTVFLTRPGYPSTVIASGNEALQILYLDNYSQGSRRQFLAGRSNVLRYL caa44873 IEMACPHGYQRVWAKLRTGSVYIVPAGHPITEIASTNSRLQILWFDLNTRGNERQFLAGKNNVLNTL caa91187 TQIVYPNGSAAATQRVSEGSVFFVPQNFPMCQIASQSGSFEFMGFTTSSRPNRPQFLAGSNSVLKGI ef136753 FOISFPNGTSALNKSVKEGTIVFVPRYFPMSOIASREGALEFVGFSTSAAPNNPOFLCGASSVLKAL efj09997 IQVSYPNGTNALDKELDESKVVFVPRYFPMCQIASRNGDFEFVGFSTSSRRNRPQFLAGSNSVFKAF caa57848 VQVVGHNGQTVLDDTVREGDLVVFPQYFAVMKRA-GNNGFEWVSFKTSASPMRS-PLAGSTSTIKGM ern18747 IQVIGTNGRKVMDDRVHEGDVFVIPQYFTAMSKA-GNEGLEWVAIKTSDLPMKS-PILGHASAIKGI caa90641 IQIVQNEGRRVFDGAVKEGQFLVIPQLHAIAKQA-GKDGLEWISFTTSDSPIRS-TLTGRNSVLKAM caa44874 IEVVGDEGRSVFDGRVREGQFIVIPQFYAVIKQA-GDEGFEWITFTTSDISFQS-FLAGRQSVLKAM efj10701 IEVVRENGEQAVEADMDEGSLLVVPANYPSAKLA-GNEGLDFAVIYKTHLPIES-YLAGRNSVYKGV efj09931 VQVARDSGEQALDEPVQEGSLVIVPANHPSAKLA-GKQGLNYYSIFTNDQPIES-YMAGRNSVYRGI gag85384 IQVAAPGGTNALDQEVSAGQLVVIPKNFPFITMAARDSPLEWVDLYMACSSLCFCFVAGNDSVYKAL gag91246 IQVAGPGGKTVLDRAVGEGDLVVIPKHFPAAKLAARAAPLEWVDVMTSPAPLSF-FLAGANSVYKAF eie22877 VGVFLEPGVF-EEAVLGAGDVGYAPKGSGHWLQNPSETEDAYILIFDDGPF----TSIELPWVLGNV eie22888 AGVFMKPGEY-EDGFLGPGDAGYAPKGSAHWLLNASNDTDAYVLIFDDGTF----TNIDLPWFIGNV ack49092 MGIFHGEGRN-NTVEFHPSDVGYVPQSIGHYVKNIGKDTLRFLEVFTSAKY----ADISLASWMNNV acb94629 MTVFANEGRA-NTFDFSPSDVGYVPRSTGHYILNTGKDTLRFLEVFTTGQY----SDISLTSWMANT aie72611 LTVFASEGKA-SVSRLQQGDVGYVPKGYGHALRNGSQKPVDILVVFNDGDY----QSIDLNSWITSN

Fig. 2.7. Alinierea secvențelor de aminoacizi ale superfamiliei de proteine de rezervă, utilizate pentru construirea arborelui evolutiv, prezentat în Figura 2.8 (pozițiile finale).

60

Așa cum era de așteptat, proteinele legumino-similare din alga verde *Klebsormidium nitens* și leguminele formează un cluster comun pe arborele evolutiv al globulinelor de rezervă (Figura 2.8). Prin urmare, proteinele legumino-similare din algele verzi *Klebsormidium nitens* reflectă trăsăturile caracteristicile celui mai vechi precursor specific al proteinelor legumino-similare din struțișor și al leguminelor din semințe [101].



#### Fig. 2.8. Calea evolutivă a globulinelor de rezervă din semințe

Regiunile secvențelor de aminoacizi, utilizate pentru analiză, corespund domeniilor C-terminale (165 poziții ale alinierii multiple, Fig.3.7a-c) ale proteinelor care reprezintă calea directă a evoluției globulinelor de rezervă de la OD monodomenică cam64834 din actinobacteria *Mycobacterium abscessus*, folosită ca rădăcină. Numerele de deasupra ramurilor reflectă suportul statistic al clusterelor (% din 1000 de replicări).

#### 2.3. Succesiunea formării structurilor globulinelor 78 și 118

Precursorul evolutiv direct al globulinelor de rezervă este proteina legumino-similară din alga verde *Klebsormidium nitens* (Figura 2.8). Așa cum am menționat mai sus, BLAST căutarea nu detectează în genomul algei verzi *Klebsormidium nitens* gene, care codifică în mod specific proteinele vicilino-similare. Astfel, în conformitate cu informațiile disponibile până în prezent, divergența în timpul evoluției globulinelor de rezervă în clustere independente de legumine și viciline a avut loc la nivelul plantelor cu spori. Structura unui precursor comun al leguminelor și vicilinelor, probabil este reflectată de proteina efj19431 din struțișor, care ocupă o poziție intermediară între proteina OD-similară din alga verde și clusterele leguminelor și vicilinelor în timpul analizei evolutive a sumei secvențelor domeniilor N- și C-terminale ale proteinelor bidomenice [2]. În acest caz, cea mai mare parte a secvenței domeniului C-terminal al acestei proteine este veridic legumino-similară [2]. Mai mult decât atât, este legumino-similară și cea mai mare parte a secvenței din domeniul C-terminal al proteinei vicilino-similare caa91187 din ferigă [4, 98]. Totalitatea acestor observații permit de a presupune că, în comparație cu vicilinele, structurile primare (precum și cele superioare) ale leguminelor reflectă trăsături mai vechi.

În conformitate cu rezultatele analizei evolutive (Figura 2.8), precursorul comun al leguminelor și vicilinelor, cu structura terțiară și cuaternară cunoscută, este OD bacteriană 2vqa. Prin urmare, structura OD 2vqa poate fi folosită ca rădăcină pentru a stabili nivelul relativ de antichitate a structurilor primare și superioare ale leguminelor și vicilinelor. Conform rezultatelor suprapunerii spațiale a structurilor terțiare ale subunităților, leguminele sunt mult mai aproape de OD 2vqa, decât vicilinele (Figura 2.9, coloanele albe). Antichitatea relativă nu numai a structurilor terțiare, ci și a structurilor primare ale leguminelor este relevată de rezultatele alinierii structurale a secvențelor de aminoacizi ale subunităților OD 2vqa și subunităților globulinelor de rezervă (Figura 2.9, coloanele negre).

La suprapunerea spațială a structurilor trimerilor OD 2vqa și globulinelor de rezervă coincid în mod satisfăcător (RMSD  $\geq$  1.6 Å) după poziția atomilor C<sup> $\alpha$ </sup> numai două (A și B) din trei subunități. Relațiile dintre structurile OD 2vqa, leguminei 3fz3 și vicilinei 1uik sunt ilustrate de rezultatele, prezentate în Figura 2.10, prin suprapunerea spațială a dimerilor lor în componența trimerului. Legumina și OD sunt similare la nivelul structurilor terțiare ale domeniilor N- și C-terminale, ale subunităților și dimerilor acestora, în timp ce asemănarea structurală dintre OD și viciline se observă aproape exclusiv la nivelul domeniilor C-terminale. Aceste informații sunt în concordanță cu rezultatele comparării structurile primare ale OD și vicilinelor (Figura 2.9).



Fig. 2.9. Compararea structurilor terțiare (coloane albe) și primare (coloane negre) ale subunităților globulinelor de rezervă și a OD 2vqa din cianobacterii Synechocystis sp. N<sub>3D</sub> – numărul atomilor C<sup>α</sup> ale resturilor de aminoacizi, care coincid după poziție la suprapunerea spațială (RMSD de la 1,3 până la 1,6 Å) a structurilor terțiare ale subunităților. N<sub>1D</sub> – numărul de resturi de aminoacizi identice în secvențele structurate ale subunităților. Legumine – 1fxz (1) și 3d5h (5) *Glycine max*; 3ksc (2) *Pisum sativum*; 2e9q (3) *Cucurbita maximă*; 3fz3 (4) *Prunus dulcis*; 3qac (6) *Amaranthus hypochondriacus*; 3c3v (7) *Arachis* hypogaea; 3kg (8) Brassica napus. Viciline – 4lej (9) Pinus koraiensis; 2cv6 (10) Vigna radiata; lipk (11) și 1uik (13) Glycine max; 2cau (12) Canavalia ensiformis; 2phl (14) Phaseolus vulgaris; 3smh (15) Arachis hypogaea

Rezultatele prezentate mai sus indică antichitatea relativă a structurilor primare, terțiare și cuaternare ale leguminelor, în timp ce trăsăturile antice ale vicilinelor s-au păstrat predominant la nivelul structurilor primare și terțiare ale domeniilor lor C-terminale.



Fig. 2.10. Diagramele panglică ale structurilor subunităților A și B în componența trimeriilor OD 2vqa din cianobacterii Synechocystis sp. (a), leguminei 3fz3 Prunus dulcis (b) și vicilinei 1uik Glycine max (c), suprapunerea spațială, dar prezentate separat Pentru claritate, structurile subunităților A și B, prezentate în figură, sunt extinse orizontal. N și C – respectiv, domeniile N- și C-terminale. Porțiunile întunecate ale diagramelor corespund atomilor C<sup>α</sup>, care nu coincid în poziție în structurile combinate ale OD/legumine (a, b) și OD/viciline (c).

#### 2.4. Controlul degradării globulinelor de rezervă în semințele germinative

O diferența semnificativă între proteinele legumino- și vicilinosimilare din *Selaginella moellendorffii* și globulinele de rezervă din spermatofite este apariția în ultimele a regiunilor nestructurate extinse în structura domeniilor [2]: buclei între  $\beta$ -strendurile E și F și între  $\beta$ -barrel și  $\alpha$ -helixuri în domeniul N-terminal al leguminelor; buclei EF în domeniul C-terminal al vicilinelor. Aceste bucle, precum și extensia N-terminală în vicilinele de tipul convicilinelor și linkerul alungit între domeniile leguminei și vicilinei formează straturile de suprafață hidrofile în structura globulinelor de rezervă [7, 29], care sunt sensibile la atacul proteolitic [15, 102, 103]. Un indicator al sensibilității potențiale la proteoliză a regiunilor din structura proteinei este prezența în ele a resturilor de aminoacizi cu accesibilitate crescută la solvent [2, 11].

Degradarea proteolitică a globulinelor de rezervă *in vitro* și în semințele germinative are loc în două etape [2]. În prima etapă se observă proteoliza limitată, fie în mod predominant, fie exclusiv, care constă în scindarea unei părți din regiunile sensibile ale moleculei, menționate mai sus. După finalizarea proteolizei limitate, se formează un produs macromolecular relativ stabil, care în continuare este scindat exclusiv prin mecanismul "una câte una"[67]. Viteza proteolizei "una câte una"este limitată de scindarea primei legături peptidice în fiecare moleculă de proteină [66, 67]. Scindările ulterioare în această moleculă au loc cu o vitează în creștere și, prin urmare, procesul se desfășoară ca o reacție de ordinul pseudo-unu [65, 66].

În funcție de densitatea de împachetare a moleculei, globulina de rezervă nativă poate fi sensibilă sau insensibilă pentru degradare prin mecanismul "una câte una"[2]. În primul caz, proteoliza limitată și "una câte una"se desfășoară simultan, independent una de cealaltă (vicilina din soia [11, 14] și legumina din ovăs [2]). În al doilea caz, proteoliza "una câte una" începe doar după finalizarea proteolizei limitate (vicilina din fasole [11], legumina din soia [9] și dovleac [10]). Astfel, în ultimele cazuri, proteoliza limitată duce la apariția în structura globulinei de rezervă a legăturei peptidice, a cărei scindare inițiază o degradare profundă a moleculei sale prin mecanismul "una câte una".

Constanta de viteză a proteolizei "una câte una"a proteinei depinde de starea conformațională a moleculei sale [104]. Aparent susceptibilitatea, menționată anterior, a globulinelor de rezervă [9, 10, 11] la proteoliza "una câte una" este dobândită în final datorită modificărilor regulate în structura terțiară a subunităților în timpul proteolizei limitate. Cu toate acestea, analiza reacțiilor succesive ale proteolizei limitate a glicininei (legumina din soia) cu papaină [9] permite de a presupune, că capacitatea de proteoliză "una câte una" a acestei proteine este inițial dobândită ca rezultat al modificărilor nu atât a structurii terțiare, cât a celei cuaternare a moleculei sale.

Proteoliza limitată a glicininei cu papaina se termină cu scindarea secvenței C-terminale extinse a α-catenelor (domeniul N-terminal), care include nu numai linkerul interdomenic

nestructurat, ci și regiunea adiacentă secvenței care formează un grup de  $\alpha$ -helixuri în proteina nativă [9] (Figura 2.11). De la o treime până la jumătate din legăturile de hidrogen, responsabile de interacțiunea dintre subunitățile A și B în homotrimerul glicininei 1fxz și 2d5h pot fi atribuite legăturilor formate cu participarea acestor helixuri [7, 9].

Astfel, în trimerii gliceninei-P (produsul final al proteolizei limitate cu papaină), interacțiunile dintre subunități sunt slăbite și, spre deosebire de glicinina nativă, glicinina-P dobândește predispoziția de a disocia. Capacitatea glicininei-P de disociere ireversibilă nu a fost detectată experimental [9]. Cu toate acestea, acest lucru nu exclude posibilitatea existenței unei disocieri reversibile, în care echilibrul de disociere să fie deplasat spre forma asociată.

Ca urmare a disocierii glicininei-P, în structura subunităților ei sunt dezgolite legăturile peptidice, potențial sensibile la atacul proteolitic, care sunt mascate în glicinina nativă, sunt expuse în structura subunităților sale. Astfel, suprafața subunităților disponibilă la solvent a glicininei-P 1fxz nedisociate (suma valorilor ASA) după disociere sa se mărește de 1,2 ori (de la 12500 până la 15400 Å<sup>2</sup>). După disocierea glicininei-P 1fxz, în cea mai mare măsură crește accesibilitatea la solvent a resturilor de aminoacizi în regiunea buclei între strendurile *I* și *J* a  $\alpha$ -catenelor. Astfel, ASA restului Leu152, care corespunde specificității de substrat a papainei [105], crește de la 40,4 la 60,7%.

Scindarea presupusă a legăturii peptidice Leu152-Glu153 în regiunea buclei IJîndepărtează secvența suplimentară C-terminală a  $\alpha$ -catenelor (Figura 2.11 a, b). Această scindare elimină restul Phe163 conservativ, care este implicat în formarea și stabilizarea trimerilor subunităților [38], în toate structurile cunoscute ale leguminelor, precum și o serie suplimentară de resturi de aminoacizi, care în glicinina nativă formează legături de hidrogen intersubunitare (Figura 2.11).

În scenariul ipotetic descris al proteolizei "una câte una" a glicininei-P, molecula de substrat disociată este îndepărtată imediat după apariția sa din amestecul de reacție. În acest caz, viteza proteoliză "una câte una" a glicininei-P este determinată de concentrația substratului disociat.





# Fig. 2.11. Glicinina 1fxz din soia: regiunile structural organizate ale α-catenei (a) și diagrama panglică a structurii ei terțiare (b)

Resturile, care formează legături de hidrogen între α-catena subunității A şi β-catena subunității B în homotrimerul glicininei [24] sunt prezentate cu majuscule (a) sau prin sfere (b); Gln155 și Ser244 formează câteva legături intersubunitare. În paranteze pătrate (a) sau linii punctate (b) este arătată poziția regiunilor nestructurale ale α-catenei: bucla *EF* (1), bucle strendul *J* (și αul h1 [6]) și un grup de α-helixuri C-terminale (2), și lincherul interdomenic (3). Porțiunea nestructurală scurtă h2' este specifică numai pentru glicinina 1fxz [6]. Resturile evisențiate cu italic (a) sau porțiunile întunecate ale diagramei (b) sunt eliminate la formarea glicininei-P (datorită scindării porțiunii centrale a sectorului 2 [73]) și după disocierea sa ca urmare a clivajului presupus al legăturii peptidice Leu152-Glu153.

#### 2.5. Concluzii la capitolul 2.

1. Ipoteza privind originea globulinelor de rezervă din semințe din OD bacteriene bidomenice (Figura 2.3), bazată pe analiza structurilor primare ale proteinelor omoloage din plante și din bacterii, este confirmată de suprapunerea spațială nu doar a structurilor secundare și terțiare ale domeniilor lor (Figura 1.9 și 1.10) și subunităților, dar și a structurilor cuaternare .

2. În genomul algele verzi au fost depistate doua tipuri de gene, care codifică proteinele bidomenice, reflectând două etape succesive ale evoluției globulinelor de rezervă. Aceste proteine OD-similare din *Coccomyxa subellipsoidea* detectate anterior [2], similare la nivelul structurilor primare și superioare cu OD bacteriene bidomenice (capitolul 1) și, așa cum s-a arătat pentru prima dată în această lucrare, proteinele legumino-similare din alga *Klebsormidium nitens*, conform tuturor trăsăturilor analizate (Figura 2.4-2.8) sunt precursorii evolutivi direcți ai proteinelor legumino-similare din plantele cu spori și, prin urmare, a leguminelor din semințe.

De remarcat că *Coccomyxa subellipsoidea* și *Klebsormidium nitens* aparțin la două filumuri diferite de alge (Chlorophyta și Charophyta, respectiv). După o serie de trăsături (biologia dezvoltării și paleobotanica), strămoșii plantelor terestre pot fi algele harofite, dar nu algele clorofite. Prezența precursorilor evolutivi specifici ai globulinelor de rezervă din semințe în genomul algei *Klebsormidium nitens*, arătate în acest studiu, reprezintă o dovadă moleculară semnificativă în favoarea acestei ipoteze.

3.  $\beta$ -barrel-ul cupinic împreună cu grupul de  $\alpha$ -helixuri, care este elementul bazal al structurii OD bacteriene, a servit ca bază pentru formarea independentă în timpul evoluției a structurilor omoloage ale subunităților globulinelor de rezervă din semințe și germinelor din plante. Pentru prima dată s-a arătat prezența în alge a proteinelor, care sunt precursorii evolutivi necondiționați ai globulinelor de rezervă. Divergența ramurii evolutive a globulinelor de rezervă în clustere independente de legumine și viciline a avut loc la nivelul unor astfel de plante vasculare cu spori, cum ar fi struțișorii și ferigile. Comparativ cu vicilinele, structurile primare și terțiare ale leguminelor cel mai mult reflectă trăsăturile străvechi ale precursorului ipotetic comun al globulinelor de rezervă.

4. Degradarea profundă a leguminei din soia în semințele germinative și *in vitro* este precedată de proteoliza limitată, care slăbește interacțiunile dintre subunități în oligomerul său. Se presupune că degradarea profundă a leguminei din soia prin mecanismul "una câte una" este inițiată de disocierea moleculelor sale.

### 3. LEGITĂȚILE PROTEOLIZEI GLOBULINEI 11S DIN SEMIȚELE DE GINKGO

S-a constatat, că proteoliză limitată a globulinelor de rezervă 11S din soia și din dovleac [9, 10] reprezintă o etapă necesară, care precedă degradarea masivă a acestor proteine (capitolul 1). Astfel, globulinele 11S din semințele acestor plante sunt protejate de degradarea profundă prematură. Există oare un rol de reglare corespunzător al proteolizei limitate a globulinelor 11S la un nivel evolutiv atât de timpuriu ca Ginkgophyta? Pentru a răspunde la această întrebare a fost realizat un studiu al mecanismului proteolizei ginnacinei, globulina de rezervă 11S din semințele de ginkgo, cu papaină și analiza tuturor structurilor cristaline cunoscute ale globulinelor 11S în comparație cu structura model al ginnacinei.

#### 3.1. Materialul și metodele de cercetare

**Izolarea și purificarea ginnacinei.** Făina din semințele de ginkgo (*Ginkgo biloba L.*) a fost extrasă cu NaCl 1 M în soluția tampon A (0,02 M Tris-HCl, pH 8,0, 0,02% NaN<sub>3</sub>, 1 mM EDTA), în raport de 1:10 (masă/volum). Ginnacina a fost precipitată din extract cu sulfat de amoniu (fracția 0-60% de saturație). Precipitatul s-a dizolvat în 3,0 M NaCl în soluția tampon A și soluția s-a aplicat pe o coloană (2,6 x 9 cm) cu fenil-Sepharose CL-4B (Pharmacia Biotech, Suedia), echilibrată cu 3,0 M NaCl în soluția tampon A. După îndepărtarea fracției neabsorbite, ginnacina a fost eluată cu NaCl 1 M în soluția tampon A. Preparatul rezultat a fost purificat suplimentar prin gel-filtrare pe o coloană (0,9 x 70 cm) cu Sephacryl S-300 (Pharmacia Biotech, Suedia), echilibrată cu NaCl 1 M în soluția tampon A.

**Proteoliza ginnacinei.** Pentru hidroliza ginnacinei, amestecul de reacție, care conținea 5 mg/ml substrat și 100 μg/ml papaină (Sigma Life Science, SUA) în soluția tampon B (tampon fosfat citrat 0,037 M, pH 5,6, NaCl 1 M, NaN<sub>3</sub> 0,02%, 1 mM EDTA, 10 mM 2-mercaptoetanol) a fost incubat timp de 24 ore la 30°C. Pentru a studia progresul hidrolizei periodic s-au luat alicote din amestecul de incubare și reacția a fost stopată prin adăugarea acidului tricloracetic (ATA) până la concentrația finală de 5% (masă/volum). În unele cazuri, reacția a fost stopată prin adăugarea E-64 (trans-epoxysuccinnyl-L-leucylamido-(4-guanidino) butan, Sigma, Life Science, SUA), până la concentrația finală de 10 μM și probele proteice rezultate au fost investigate prin gel-filtrare pe o coloană cu Sephacryl S-300, urmată de analiza lor prin electroforeză.

**Electroforeza și analiza electroforegramelor.** Electroforeza în gel de poliacrilamidă a proteinei, disociată cu dodecilsulfat de sodiu (SDS-PAGE) a fost efectuată într-un gel de 15% (g/v)

în sistemul tampon Laemmli [106]. Înainte de electroforeză, ATA rezidual în probele proteice precipitate a fost îndepărtat prin spălare cu acetonă, precipitatele au fost uscate și au fost dizolvate în condiții standard într-o soluție tampon, care conține sau nu conține 2-mercaptoetanol. Au fost utilizați markerii moleculari Bio-Rad (SUA). Electroforegramele au fost scanate (ImageScanner III, GE Healthcare) și au fost analizate, utilizând programul Phoretix 1D Gel Analysis v.5.10. Cantitățile molare relative ale polipeptidelor, detectate prin electroforeză în ginnacina intactă și în proteina reziduală după proteoliză, s-au calculat în funcție de masele lor moleculare și de concentrațiile de masă relative.

Analiza cineticii proteolizei. Pentru a analiza cinetica proteolizei, proteina reziduală în hidrolizate precipitată cu ATA a fost determinată prin legarea colorantului [107]. Masa moleculară medie numerică a subunităților proteinei inițiale și a restului său cu masa moleculară mare în hidrolizate au fost calculate din datele masei moleculare a polipeptidelor, detectate prin electroforeză, și conținutul lor de masă relativ. Concentrația proteinei reziduale în hidrolizate P<sup>t</sup> și masa sa moleculară medie numerică M<sup>t</sup> au fost exprimate în % din valorile inițiale corespunzătoare  $P^0$  și  $M^0$ .

Analiza cineticii proteolizei a fost efectuată urmând strategia descrisă anterior [9, 10]. În timpul proteolizei după mecanismul mixt, când proteoliza limitată și "una câte una" decurg fie paralel, independent una față de alta, fie succesiv, contribuția proteolizei limitate la scăderea generală a concentrației de masă a proteinei în hidrolizat ( $P\%^0 \rightarrow P\%^t$ ) este egală cu  $M\%^0$ - $M\%^t$ . Prin urmare, ecuația  $P\%^t + (M\%^{0-}M\%^t) = f(t)$  descrie exclusiv cinetica proteolizei "una câte una", care decurge ca reacție de ordinul pseudo-unu [8, 38]. Respectiv, modificarea în timpul reacției a valorii relative a masei moleculare a substratului proteic ( $M\%^0 \rightarrow M\%^t$ ) descrie exclusiv cinetica proteolizei limitate.

Modelarea structurilor proteice și metodele de analiză computerizată. Modelarea omoloagă a structurilor terțiare și cuaternare a ginnacinei a fost efectuată cu ajutorul programului Swiss-Model [91, 92], utilizând în calitate de matriță structura cristalină pdb|3qac a globulinei 11S din amarant *Amaranthus hypochondriacus*. Acest program de asemenea a fost utilizat pentru modelarea structurilor cuaternare ale globulinelor 11S, pentru care au fost publicate structura terțiară a numai unei subunități A (pdb|2e9q, pdb|3c3v, pdb|3qac). Pentru a modela structura terțiară a ginnacinei, au fost utilizate de asemenea două metode alternative: programul SSPRO 8 (Scratch Protein Predictor) pentru prezicerea structurilor secundare și programul DisEMBL 1.5 pentru prezicerea regiunilor nestructurate.

Suprafața accesibilă a resturilor de aminoacizi (ASA) la solvent în structurile cristaline și model ale oligomerilor globulinelor 11S, exprimată în Å<sup>2</sup>, a fost calculată cu ajutorul programul <u>http://cib.cf.ocha.ac.jp/bitool/ASA/</u>. Valoarea relativă ASA a restului X în globulina 11S este exprimată în % din accesibilitatea lui în tripeptida GXG. Programul DeepView/Swiss-PdbViewer v.3.7 a fost utilizat pentru a construi diagramele panglică și pentru a identifica legăturile de hidrogen intersubunitare în oligomerii globulinelor 11S, și pentru determinarea deviației pătratice medii (root-mean-square deviation, RMSD) a poziției atomilor C<sup> $\alpha$ </sup> la suprapunerea spațială a structurilor terțiare ale subunităților ginnacinei (model) și altor globuline 11S (structurile cristaline).

Identificarea punctelor probabile de scindare a ginnacinei în timpul proteolizei a fost efectuată utilizând rezultatele electroforezei (compoziția fragmentelor de polipeptidă și masele lor moleculare aparente conform SDS-PAGE, transformarea succesivă a fragmentelor mari în cele mai mici, detectate prin schimbările corespunzătoare în intensitatea zonelor lor) și în conformitate cu informațiile de mai jos.

1. Rezultatele comparării structurilor terțiare ale ginnacinei și ale altor globuline 11S, a căror proteoliză limitată este descrisă [96], permit *a priori* determinarea poziției regiunilor cele mai probabile de scindare a secvenței de aminoacizi a ginnacinei. Accesibilitatea sporită la solvent a acestor regiuni în structurile ginnacinei și a altor globuline 11S poate fi considerată ca dovadă a sensibilității lor potențiale la proteoliza limitată [2].

2. Corespunderea punctelor presupuse de scindare din regiunile potențial sensibile ale moleculei ginnacinei cu specificitatea primară de substrat a papainei [105] (resturile Arg, Lys, Leu și Gly în poziția P1 în absența restului Val în poziția P1').

#### 3.2. Izolarea și purificarea ginnacinei.

În timpul salifierii partea principală a ginnacinei este precipitată la o concentrație de sulfat de amoniu de 40-60% saturație (Figura 3.1). Fracția respectivă a fost utilizată pentru purificarea ulterioară.

La cromatografia ulterioară pe coloana cu fenil-Sepharose CL-4B (Figura 3.2), după îndepărtarea fracției neadsorbite, ginnacina în cantități semnificative și mai pură este eluată la concentrația NaCl 1 M (Figura 3.3).



Fig. 3.1. SDS-PAGE proteinelor din extractul brut din semințele de ginkgo (pistele 1 și 5) și fracțiilor precipitate cu sulfat de amoniu în intervalul de concentrație (în % saturație): 0-40% (pistele 2 și 6), 40-60% (pistele 3 și 7) și 60-80% (pistele 4 și 8) în prezența (pistele 1-4) și în absența (pistele 5-8) 2-ME.



Fig. 3.2. Cromatografia pe coloana cu fenil-Sepharose CL-4B a ginnacinei, parțial purificată prin salifiere

Linia orizontală indică regiunea de eluare, ce corespunde ginnacinei purificate (Fig. 3.3. pistele 2 și 5). Fracția neadsorbită nu conține proteină și, probabil, corespunde acizilor nucleici (maximul de absorbție la 260 nm).
Cu toate acestea, după cum reiese din rezultatele electroforezei în absența 2mercaptoetanolului (Figura 3.3, linia (pista) 5) în preparatul, purificat prin cromatografie, mai sunt prezente impurități. După gel-filtrarea pe coloana cu sefacryl S-300, aceste impurități sunt complet îndepărtate (Figura 3.4, pistele 1 și 8).



**Fig. 3.3. SDS – PAGE preparatului de globulină 11S, parțial purificat prin salifiere cu sulfat de amoniu** (pistele 1 și 4) **și fracțiilor eluate la cromatografia fenil-Sepharose** cu NaCl 1M (pistele 2 și 5) și apă (pistele 3 și 6) în prezența (pistele 1-3) și în absența (pistele 4-6) 2-ME.

### 3.3. Proteoliza și cinetica proteolizei ginnacinei.

**Proteoliza ginnacinei**. Electroforeza ginnacinei intacte (Figura 3.4, pista 1) relevă două benzi ale  $\alpha$ -catenelor (28,3 kDa și 26,7 kDa) și o bandă a  $\beta$ -catenelor (21,1 kDa). Astfel, în componența moleculei hexamerice a ginnacinei [108] intră, cel puțin, două tipuri de  $\alpha\beta$ -subunități. La fel ca în celelalte globuline 11S,  $\alpha$ - și  $\beta$ -subunitățile ginnacinei sunt legate printr-o legătură disulfidică (pistele 1 și 8).

Este cunoscută secvența de aminoacizi numai a unei subunități a ginnacinei (caa90641 cu masele moleculare ale  $\alpha$ - și  $\beta$ -catenelor de 27,5 kDa și, respectiv, 21,1 kDa) [109].

Proteoliza ginnacinei începe cu formarea, cel puțin a două fragmente intermediare F1 ale  $\alpha$ -catenelor cu masele moleculare 25,8 și 25,2 kDa (Figura 3.4, pistele 2-5). La proteoliza ulterioară fragmentele F1 dispar și apar două fragmente: un fragment F3 principal cu masa moleculară 21,1 kDa și un fragment F2 minor cu masa moleculară 22,0 kDa. Fragmentul F3 coincide după mobilitate cu  $\beta$ -catenele. Remarcăm, că pe electroforegramă sunt absente produsele

cu masa moleculară mică de 20 kDa [110]. Prin urmare, sectoarele scindate ale  $\alpha$ -catenei în produsul final al proteolizei sunt scindate până la peptide cu masa moleculară mică.

Astfel,  $\beta$ -catenele ginnacinei sunt insensibile la proteoliza limitată, așa cum este tipic pentru majoritatea globulinelor 11S din semințe studiate [2], iar hidroliza  $\alpha$ -catenelor constă în scurtarea lor succesivă N- sau C-terminală.



**Fig. 3.4. SDS-PAGE ginnacinei intacte și produselor ei de proteoliză** în prezența (pistele 1-7) și în absența (pistele 8, 9) 2-ME. Pe partea stângă sunt arătate masele moleculare ale markerilor. Linia de jos indică timpul de reacție (h).

În timpul gel-filtrării pe coloana cu Sephacryl S-300 a produselor de hidroliză ale ginnacinei, formate după 24 ore de reacție (Figura 3.5), s-a observat un singur vârf al proteinei cu masa moleculară mare, care conform compoziției electroforetice corespunde probei, prezentată în Figura 3.4, pista 7.



Fig. 3.5. Gel-filtrarea pe coloana cu Sephacryl S-300 a produselor de hidroliză a ginnacinei după 24 ore de reacție.

**Cinetica proteolizei ginnacinei**. Pentru a descrie mecanismul proteolizei ginnacinei cu papaină, am studiat cinetica acestui proces bazată pe determinarea a două valori relative - concentrația de masă a proteinei P% și masa moleculară a moleculei reziduale nedisociate M%. Compararea acestor valori ne permite să determinăm care dintre cele două mecanisme ale proteolizei se observă pe măsură ce se dezvoltă acest proces - proteoliză limitată sau proteoliză "una cîte una" explozivă.

În prima etapă a proteolizei (până la 2 ore), scăderea relativă a concentrației de masă a proteinei P% în hidrolizat este egală cu scăderea masei sale moleculare M% (Figura 3.6, graficele 1 și 2). Cu alte cuvinte, în prima etapă, valorile P% și M% coincid. Aceasta indică absența în prima etapă a semnelor proteolizei "una cîte una".





I – concentrația de masă P a proteinei reziduale (precipitate cu TCA); 2 – masa sa moleculară medie M. P şi M sunt exprimate în % din valorile lor inițiale. Graficul 3 (P%+[100%-M%]) descrie cinetica exclusiv a proteolizei "una câte una". Scala ordonatei este logaritmică.

În a doua etapă (2-7 h), proteoliza limitată este completă, adică masa moleculară M% este constantă. Cu toate acestea, concentrația de masă a proteinei P% scade liniar, ceea ce este un semn al proteolizei "una cîte una". Scăderea concentrației de masă relativă a proteinei ca urmare a proteolizei limitate este egală cu 100%-M%. Prin urmare, cinetica exclusiv a proteolizei "una câte una" poate fi descrisă de al treilea grafic, care reprezntă suma valorile P% și 100%-M% obținute

experimental. Extrapolarea porțiunii liniare a celui de-al treilea grafic până la zero duce la o valoare care depăşeşte 100%.

Acest fapt indică dependența constantei de viteză a proteolizei "una câte una" a ginnacinei de modificările structurale ale acesteia, ca urmare a proteolizei limitate [111]. Astfel, ginnacina nativă, la fel ca și globulinele 11S din soia și din dovleac [9, 10], este insensibilă la proteoliza "una câte una", care începe doar după atingerea unui anumit nivel al proteolizei limitate.

# 3.4. Modelarea omoloagă a structurilor terțiare și cuaternare ale ginnacinei.

În prezent, există doar informații inițiale privind structura cristalină a ginnacinei [108]. Prin urmare, modificările structurale ale ginnacinei în timpul proteolizei limitate pot fi estimate numai pe baza modelului de structură terțiară și cuaternară și, de asemenea, pe baza informațiilor disponibile privind legitățile proteolizei altor globuline 11S din semințe.

Un punct important în modelarea structurii ginnacinei este alegerea globulinei 11S cu structură cunoscută, cea mai apropiată de ginnacină după secvența de aminoacizi. Analiza relațiilor filogenetice dintre ginnacină și globulinele 11S, ale căror structuri cristaline sunt stabilite, arată că structura globulinei 11S din amarant *Amaranthus hypoandriacus* este cea mai potrivită pentru a fi utilizată în calitate de matriță (Figura 3.7).



# Fig. 3.7. Relațiile filogenetice dintre ginnacină și globulinele 11S cu structura terțiară cunoscută

Numerele de deasupra ramurilor reflectă suportul statistic al clusterelor (% din 1000 replicări). În calitate de rădăcină, a fost utilizată secvența proteinei leguminosimilare bidomenică din struțișor *Selaginella moellendorffii*. Globulina 11S din magnolie *Magnolia salicifolia* ocupă o poziție intermediară între ginnacină și globulinele 11S cu structura cunoscută.

Analiză comparativă detaliată a structurilor primare ale ginnacinei și altor globuline 11S cu structură cunoscută, furnizată de programul Swiss Model (Tabelul 3.1) arată că trei globuline 11S (3qac, 3fz3 și 3ehk) sunt aproape la fel de potrivite ca matriță pentru modelarea structurii ginnacinei. Utilizarea primei dintre aceste proteine ca matriță este cea mai preferată în concordanță cu rezultatele analizei filogenetice prezentate în Figura 3.7.

Şablonul	Identitate,SubstituțiiAcoperire %%conservative, %		Sursa						
Globuline 11S									
3qac	44	41	97	Amaranthus					
3fz3	44	42	96	Davage					
3ehk	44	42	96	Prunus					
3kgl	40	39	95	Brassica					
2e9q	40	40	97	<i>C</i> 1:					
2evx	40	40	97	Cucurbita					
10d5	40	40	94						
2d5h	39	40	94						
2d5f	39	40	94						
1fxz	40	40	95						
lucx	41	41	95	Glycine					
1ud1	41	41	95						
1ud1	39	40	95						
1ucx	39	40	95						
3ksc	38	39	96	Pisum					
3c3v	38	40	96	Arachis					
		Globulina 7S							
4lej	27	35	81	Pinus					
		OD bacterian	ă						
2vqa	18	29	75	Synechocystis					

# Tabelul 3.1. Analiza oportunității utilizării ca model structurile cristaline ale globulinelor 11S pentru modelarea structurii terțiare și cuaternare a ginnacinei.

Notă: Acoperirea este numărul relativ de resturi de aminoacizi (în % din secvența totală a globulinei 11S) utilizați pentru analiză.



## Fig. 3.8. Diagramele panglică ale structurilor α-catenelor ginnacinei (a) şi globulinei 11S pdb|3qac din amarant Amaranthus hypochondriacus (b) suprapuse spațial, dar prezintate separat

Abaterea standard (RMSD) la alinierea a 368 atomii C<sup> $\alpha$ </sup> (suma secvențelor  $\alpha$ - și  $\beta$ -catenelor ale acestor proteine) a fost 0,32Å. La baza structurii model a  $\alpha$ - și  $\beta$ -catenelor ginnacinei (ca și la alte globuline 11S [38]) stă  $\beta$ -barrel-ul format din  $\beta$ -strendurile antiparalele *BCDEFGHI*, cuplat cu un grup de  $\alpha$ -helixuri *h1*, *h2*, *h3* și suplimentat cu  $\beta$ -strendurile *A*, *A'* și *J*, *J'* și  $\alpha$ -helixul *h0*;  $\beta$ strendurile *E'* și *F'*, caracteristice numai pentru  $\alpha$ -catene, sunt prezente în structurile multor globuline 11S [38]. Zonele potențial sensibile la proteoliza limitată în structurile cristaline ale globulinelor 11S [2]:  $\alpha$  – buclă dintre  $\beta$ -strendurile *E* și *F*; *b* – bucla dintre  $\beta$ -barrel și  $\alpha$ -helixuri; *c* – regiunea C-terminală. Asteriscurile indică pozițiile resturilor de aminoacizi în structura ginnacinei (a) cu accesibilitate sporită (> 90%) la solvent, care corespund specificității de substrat a papainei [105]. Partea întunecată a diagramei corespunde regiuni C-terminală, eliminate, probabil, la sfârșitul proteolizei limitate.

		A' A hO B
G	9	QREQLQQSCRFDRLNAQEPTQRITSEGGSVELLNVEDSEQFQCAGVAPLRETLNPNA
А	6	refqQGNECQIDRLTALEPTNRIQAERGLTEVWDS-NEQEFRCAGVSVIRRTIEPHG
		· * · * · * * * * * * * * * * · · · · ·
		C D E E'
G	66	LSLPRYTNTPTMAYVVEGEGRLGVVFPGCPETFQSSTS <b>R</b> G
А	62	LLLPSFTSAPELIYIEQGNGITGMMIPGCPETYesqsqqfqqqederireqqsrkfq
		<u>* ** :* :* : *: :* *: :* ** *: :</u> ******:* · · · : *
		F' F G H I
G	106	GEGQQSQ <u>ERSQKIRRVRRGDVVAI</u> PAGV <u>AYWLYN</u> DGNRRLQIVAIADTSNHQNQLDQ
А	119	mrgdrfQDQHQKIRHLREGDIFAMPAGVSHWAYNNGDQPLVAVILIDTANHANQLDK
		.*:: *:: *** <del>*</del> :*************************
		J h1 h2 h3
G	163	TYRP-FYLAGSAPSGAQKAAGATSIGDNILQGFDTDTLAEAMGISQDTAR
А	176	NFPTR <u>FYL</u> AGKPQQEHSgehqfsresrrgerNTGNI <u>FR</u> GFE <u>TRLLAESF</u> GVS <u>EEIAQ</u>
		·: ****· · · · · · · · · · · · · · · · ·
		- /
_		J'
G	212	RIQQN-Q <b>K</b> KGLIVKVERGLRMPGPPSDDYEREREREGN↓NVEEFYCSMRLR
А	233	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA
A	233	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::*       ::*         *       ****         *       ****         *       ****
A	233	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::*       ::*         ::*       :*         ::*       :*         ::*       :*         ::*       ::*         ::*       ::*         ::*       :*         ::*       :*         ::*       :*         ::*       :*
A	233	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn $\downarrow$ gveeTICSARLA::* : :.:* **:*: ***: ***: ***: ***:****ZA'Ah0BCHNADDSEDADVYVPNGCELNTVNELKLPALESLELGAERCILOPNAMEAPSW-LNAH
A G	233 260 280	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*:: *** :*         Z       A'         A       h0       B         HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDDSKADVYTDEACDLTTTVNSENL DILDULDU SAAKCYLYDNAMMADHYNLNAH
A G A	233 260 289	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::*:       ::*:         ::*:       ::*:         Z       A'         A       h0       B         C       HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         ****       ************************************
A G A	233 260 289	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*:: *** ::**         Z       A'         A       h0       B         C       HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         *.**      ****
A G A	233 260 289	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*: **: **: **: **: **         Z       A'         A       h0       B         C       HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         *.*****: **       ******         D       E       F       G       H       I
A G A G	233 260 289 316	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*:: ***: ***: ***         Z       A'         A       h0       B         C       HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         *.******       ::** :** *** :***         D       E       F       G       H         AVMYVTRGOGRIOIVONEGRRVFDGAVKEGOFLVIPOLHAIAKOAGKDGLEWISFTT
A G A G A	233 260 289 316 346	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*:: ***: ***: ***         Z       A'         A       h0       B         C         HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         *.**         D       E         F       G         H         AVMYVTRGQGRIQIVQNEGRRVFDGAVKEGQFLVIPQLHAIAKQAGKDGLEWISFTT         NIMYCVRGRGRIOIVNDOGOSVFDEELSRGOLVVVPONFAIVKOAFEDGFEWVSFKT
A G A G A	233 260 289 316 346	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*: **: **: **: **: **: **:
A G A G A	233 260 289 316 346	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : :::* **::*:: ***: ***: ***: ***:
A G A G A	233 260 289 316 346	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn↓gveeTICSARLA         ::* : ::* **:*:*:**: *** ::*         Z       A'         A       h0       B         C       HNADDSEDADVYVRNGGRLNTVNRLKLPALRSLRLGAERGILQPNAMFAPSW-LNAH         VNVDDPSKADVYTPEAGRLTTVNSFNLPILRHLRLSAAKGVLYRNAMMAPHYNLNAH         *.*****:::**         **
A G A G A G	233 260 289 316 346 373	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpngveeTICSARLA $KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpngveeTICSARLAKLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpngveeTICSARLAX* * * * * * * * * * * * * * * * * * * $
A G A G A G A	233 260 289 316 346 373 403	KLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn $\downarrow$ gveeTICSARLAKLQAEQDDRGNIVRVQEGLHvikppsraweereqgsrgsrylpn $\downarrow$ gveeTICSARLA::* : :::* **:*:*:********************
A G A G A G A	233 260 289 316 346 373 403	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

# Fig. 3.9. Alinierea secvențelor de aminoacizi ale α-catenelor ginnacinei (G) și globulinei 11S din amarant *Amaranthus hypochondriacus* (A)

Literele mici arată regiunile nestructurate în structura cristalina pdb|3qac *Amaranthus hypochondriacus*. Săgeata indică hotarul dintre α- și β-catenele (domeniile N- și C-terminale) globulinelor 11S. Resturile cu accesibilitate sporită (valoare relativă a ASA>90%) la solvent, care corespund specificității de substrat a papainei sunt indicate cu bold [105]. În modelul structural al subunităților ginnacinei sunt prezente toate structurile secundare, care sunt caracteristice nu doar globulinei 11S din amarant *Amaranthus hypochondriacus* (structura ei a fost utilizată ca model pentru construirea modelului ginnacinei, Figura 3.8 și 3.9), dar și pentru globulinele 11S din soia și din dovleac (Figura 3.10), precum și toate celelalte structuri cristaline cunoscute ale globulinelor 11S [38]. Acest lucru demonstrează veridicitatea structurii terțiare model a ginacinei.

		A' A h0 B
2EVX	19	hryqsprACRLENLRAQDPVRRAEAEAGFTEVWDQDN-DEFQCAGVNMIRHTIR
10D5	1	itsskfNECQLNNLNALEPDHRVESEGGLIETWNSQH-PELQCAGVTVSKRTLN
GINb	9	QREQLQQSCRFDRLNAQEPTQRITSEGGSVELLNVEDSEQFQCAGVAPLRETLN
		· · * · · · · * · * · * · * · · · · · ·
		C D E E'
2EVX	72	PKGLLLPGFSNAPKLIFVAQGFGIRGIAIPGCAETYQTDLRrsqsagsaF
10D5	54	RNGLHLPSYSPYPQMIIVVQGKGAIGFAFPGCPETFEKPqqqssrrgsrsqqql
GINb	63	PNALSLPRYTNTPTMAYVVEGEGRLGVVFPGCPETFQSSTSRGGEGQQS
		· · · · · · · · · · · · · · · · · · ·
		F'FGHI
2EVX	122	KDQHQKIRPFREGDLLVVPAGVSHWMYNRGQSDLVLIVFADTRNVANQIDPYLR
10D5	108	QDSHQKIRHFNEGDVLVIPPGVPYWTYNTGDEPVVAISLLDTSNFNNQLDQNPR
GINb	112	QERSQKIRRVRRGDVVAIPAGVAYWLYNDGNRRLQIVAIADTSNHQNQLDQTYR
		: *** * **: * ** ** *: : : : ** * ** *
		T b1 b2
	176	
	162	
CIMP	166	
GIND	TOO	PFILAGSAPSGAQKAAGAISIGDNILQGFDIDILAE
		•••••••
		h3 J'
2EVX	221	<u>AFQIDGGLVRKLK</u> GEDDERD <u>RIVQV</u> DEDFEVLLpekdeeersrgryi[ 7]N↓
10D5	216	SFNTNEDTAEKLRSPDDERKQIVTVEGGLSVISPKWqeqedededed[57]N
GINb	202	$\underline{AM}GISQDTARRIQQN-QKKGLIVKVERGLRMPGPPSDDYEREREREG[ 0]N \downarrow$
		::::: : * * :: * . : : * :: * ::

# Fig. 3.10. Alinierea secvențelor de aminoacizi ale α-catenelor globulinelor 11S din dovleac 2evx, din soia 10d5 și ginnacinei (GINbi)

Numerele din parantezele pătrate corespund numărului resturilor de aminoacizi lipsă din cele trei regiuni extinse ale secvențelor globulinelor 11S [2]: buclele dintre β-strendurile *E'* și *F'*, dintre β-strendul J și α-helixuri și regiunea C-terminală. Săgețile arată poziția punctelor de clivare stabilite pentru scindarea α-catenelor globulinelor 11S din dovleac [10] și din soia [9].

Metode alternative de modelare a structurii terțiare a subunităților ginnacinei. Secvențele de aminoacizi ale ginnacinei și ale altor globuline 11S (în primul rând, globulina 11S din amarant *Amaranthus hypochondriacus*, figura 3.9) sunt cu siguranță omoloage (tabelul 3.1 și figura 3.10). Aceasta oferă motive să credem că rezultatele modelării omoloage cu un grad înalt de fiabilitate reflectă principalele trăsături ale structurii terțiare reale a ginnacinei. Cu toate acestea, există abordări alternative de modelare a structurilor secundare și regiunilor nestructurate caracteristice regiunilor respective ale structurilor cristaline ale multor altor proteine, care constau în estimarea probabilității prezenței în moleculă unei anumite proteine, a cărei analogi structurali sunt necunoscuți. Utilizarea acestor metode face posibilă evaluarea gradului de fiabilitate a rezultatelor modelării omoloage a structurii terțiare a ginnacinei.



# Fig. 3.11. Compararea rezultatelor modelării structurilor secundare ale α-catenelor ginnacinei:

cu bold, modelare omoloagă; S – β-strendurile și H – α-helixurile, prezise de programul SSPRO
 8 (cu bold sunt evidențiate porțiunile care coincid). Simbolurile > arată regiunile neordonate, prezise de programul DisEMBL. Simbolurile < arată regiunile neordonate reale, prezente în structura cristalină pdb|3ac a globulinei 11S din *Amaranthus hypochondriacus*.

În general, rezultatele modelării structurilor secundare și regiunilor nestructurate în molecula de ginnacină prin metode alternative, SSPRO 8 și DisEMBL, respectiv (Figura 3.11), nu contrazic rezultatele modelării omoloage (Figura 3.9). Mai mult, 71% și 88%, respectiv în limitele  $\beta$ -strendurilor și  $\alpha$ -helixurilor în modelul omolog al ginnacinei se găsesc în structurile secundare

prognozate de programul SSPRO 8. Acesta este un argument suplimentar în favoarea metodei de modelare omoloagă.

Cu toate acestea, deoarece nu există o coincidență absolută a structurilor secundare prezise prin modelarea omoloagă și prin metodele alternative, pare rezonabilă compararea rezultatelor determinării structurilor secundare prezente real în structura cristalină pdb|3qac a globulinei 11S din amarant *Amaranthus hypochondriacus* cu rezultatele modelării structurilor secundare ale acestei proteine cu ajutorul programului SSPRO 8. Rezultatele obținute coincid incomplet (Figura 3.12) în același mod ca cel observat mai sus (Figura 3.11).



#### Fig. 3.12. Compararea structurilor secundare,

prezente real în structura cristalină pdb|3qac a globulinei 11S din amarant *Amaranthus hypochondriacus* (S – β-strendurile, H – α-helixurile) cu rezultatele modelării structurilor secundare ale acestei proteine, utilizând programul SSPRO 8. Cu bold sunt evidențiate regiunile care coincid. Literele mici corespund regiunilor nestructurate din structura cristalină pdb|3qac.

O altă dovadă în favoarea metodei de modelare omoloagă sunt rezultatele determinării deviației standard (RMSD) a atomilor C<sup> $\alpha$ </sup> la suprapunerea spațială a structurilor terțiare ale subunităților ginnacinei și altor globuline 11S. În mod natural, valoarea RMSD cea mai mică (0,32 Å pentru 368 atomi C<sup> $\alpha$ </sup> suprapuși) a fost obținută prin compararea structurii model al ginnacinei cu structura matriței, globulina 11S din amarant *Amaranthus hypochondriacus* (Figura 3.8). La suprapunerea structurii ginnacinei și tuturor celorlalte globuline 11S cu structura terțiară cunoscută în pereche (Tabelul 3.2.) valorile medii ale RMSD sunt destul de scăzute (0,98±0,07 Å pentru cei 352±6 atomi C<sup> $\alpha$ </sup> suprapuși).

Legitățile proteolizei limitate a ginnacinei se datorează structurii sale. Trei regiuni extinse în  $\alpha$ -catenele globulinelor 11S sunt potențial sensibile la atacul proteolitic [2]: a – între  $\beta$ -strendurile E' și F'; b – între  $\beta$ -barrel și grupul de  $\alpha$ -helixuri; c – regiunea C-terminală (Figura

3.8b). Fiecare dintre regiunile *a*, *b* și *c* în structura cuaternară model a ginacinei conține resturi de aminoacizi cu accesibilitate sporită la solvent (Figura 3.9), care corespund specificității de substrat a papainei [105].

Bucla dintre  $\beta$ -strendurile E' și F' în structura ginnacinei este insensibilă la acțiunea papainei, deoarece în produsele de hidroliză a ginacinei (Figura 3.4) sunt absente fragmentele cu masa moleculară mai mică de 20 kDa, care corespund jumătății N-terminale a  $\beta$ -barrel-ui ( $\beta$ strendurile A'-E', Figura 3.10). În mod similar, această buclă este insensibilă la atacul papainei în globulinele 11S din soia [9] și din dovleac [10].

Proteoliza globulinelor 11S [2, 15, 43, 50, 102], inclusiv globulinelor 11S din soia și din dovleac [9, 10] începe cu scindarea regiunii C-terminale a  $\alpha$ -catenelor. Probabil, ginnacina nu este o excepție de la această regulă, așa cum reiese din apariția fragmentelor intermediare F1 în stadiile inițiale ale proteolizei (Figura 3.4).

S-a arătat, că proteoliza ulterioară a  $\alpha$ -catenelor globulinelor 11S din soia și din dovleac cu papaină duce la scindarea buclei extinse între  $\beta$ -strendul *J* și  $\alpha$ -helixurile *h1-h3* [9, 10] (Figura 3.10). În structura model a ginnacinei această buclă conține restul Gly182 cu accesibilitate sporită la solvent și, prin urmare, cu sensibilitate crescută la atacul papainei (Figura 3.9). Clivajul legăturii peptidice Gly182-Ala183 în timpul hidrolizei ginnacinei cu papaină este foarte probabil, deoarece masa moleculară calculată a fragmentului Arg1-Gly182 (20,2 kDa) este aproape de valoarea obținută experimental (masa moleculară aparentă a fragmentului final F3 21,1 kDa, Figura 3.4).

În timpul hidrolizei ginnacinei cu papaină, precum și a globulinelor 11S din soia și din dovleac [9, 10], regiunea scindabilă dintre  $\alpha$ -helixurile *h1-h3* și  $\beta$ -strendul *J'* este distrusă cu formarea produselor cu masa moleculară mică, care nu pot fi detectate prin electroforeză (Figura 3.4). Acest lucru duce la schimbări semnificative în structură, datorită cărora ginnacina obține sensibilitate la proteoliza "una câte una". Astfel, proteoliza limitată a ginacinei, precum și a globulinelor 11S din soia și din dovleac [9, 10] joacă un rol de reglare, protejând aceste proteine de degradarea masivă prematură.

Rolul de reglare al proteolizei limitate s-a format în stadiile timpurii ale evoluției globulinelor 11S din semințe. Secvențele de aminoacizi ale globulinelor 11S din semințe din afara regiunile variabile a, b și c (Figura 3.8 și 3.9) sunt foarte conservative, inclusiv în regiunea structurală a  $\alpha$ -helixurilor h1-h3 și a  $\beta$ -strendului J' [28]. Eventual, această regiune este capabilă de degradare în timpul proteolizei nu numai a globulinelor 11S din ginkgo, soia și dovleac, dar și a altor proteine din această familie. Pentru a verifica această ipoteză, s-au comparat secvențele de aminoacizi ale  $\alpha$ -catenelor ginnacinei și tuturor globulinelor 11S cu structura terțiara cunoscută. Așa cum rezultă din Tabelul 3.2, accesibilitatea la solvent a regiunii  $\alpha$ -helixurilor *h1-h3* și  $\beta$ -strendului *J'* este de 2,3-3,4 ori mai mare decât cea a jumătății N-terminale a  $\beta$ -barrelui ( $\beta$ -strendurile *B-E*), care este insensibilă la proteoliza limitată [2]. Acest fapt evidențiază capacitatea potențială a acestei regiuni de a scinda profund în timpul proteolizei globulinelor 11S.

		Regiunea structurii							
Specia	pdb	AS	$A, Å^2$	H-legături					
		B-E	hl - J'	A'-J	hl - J'				
Ginkgo biloba		15.3	42.4	10	11				
Amaranthus hypochondriacus	3qac	11.6	39.7	14	8				
Prunus dulcis	3fz3*	13.7	33.2	16	10				
Brassica napus	3kgl*	13.5	36.8	15	12				
Cucurbita maxima	2e9q	12.9	33.9	11	11				
Glycine max	2d5h*	13.4	34.7	9	12				
Glycine max	1fxz*	15.2	34.8	11	9				
Pisum sativum	3ksc*	13.0	37.6	14	9				
Arachis hypogaea	3c3v	15.9	37.0	16	10				
Valo	13.8	36.7	13	10					

Tabelul 3.2. Accesibilitatea la solvent și legăturile de hidrogen intersubunitareîn oligomerii globulinelor 11S.

Notă: ASA – suprafața accesibilă la solvent, în medie, a unui rest de aminoacid în jumătatea Nterminală a  $\beta$ -barrelui ( $\beta$ -strendurile *B-E*) și în regiunea  $\alpha$ -helixurilor *h1*, *h2*, *h3* și  $\beta$ -strendului *J'* (*h1-J'*). H-legături – numărul legăturilor de hidrogen, formate cu participarea resturilor de aminoacizi din regiunea  $\beta$ -strendurilor *A'-J* și din regiunea C-terminală ( $\alpha$ -helixul *h1*- $\beta$ -strendul *J'*). Asterixul indică globulinele 11S cu structură cuaternară stabilită experimental.

Identificarea legăturilor de hidrogen intersubunitare în structura cuaternară a globulinelor 11S (Tabelul 3.2), realizată în acest studiu și anterior [7, 9], arată că până la jumătate din aceste legături se formează cu participarea resturilor de aminoacizi din interiorul regiunii C-terminale a subunității A din afara regiunii β-barrel-β-strendul *J*.

Se pare, că aceste legături sunt capabile de distrugere în timpul proteolizei nu numai ginacinei (Figura 3.11) și globulinelor 11S din soia și din dovleac [9, 10], dar și a altor globuline 11S cu

structura cuaternară cunoscută din cauza distrugerii regiunii  $\alpha$ -helixul *h1*- $\beta$ -strendul *J'* (Figura 3.12). Ca urmare interacțiunile intersubunitare devin semnificativ slăbite, iar globulinele 11S dobândesc capacitatea unei disocieri reversibile, care expun legăturile peptidice, potențial sensibile la proteoliză, dar mascate în moleculele native [96].

Rolul specific al restului de aminoacid Gln155 în formarea interacțiunilor intersubunitare în structura globulinelor 11S. Atrage atenția restul Gln155 din partea C-terminală a  $\beta$ -barrelui între  $\beta$ -strendurile *I* și *J* ale subunității A, care formează până la cinci legături de hidrogen (Figura 3.13 și 3.14) cu aminoacizii din regiunea  $\beta$ -barrelui al subunității B (Figura 3.15).



# Fig. 3.13. Legăturile de hidrogen (linii punctate) responsabile de interacțiunea dintre subunitățile A și B în structura cuaternară model a ginnacinei

Diagrama panglică este limitată la regiunea C-terminală a α-catenei subunității A, eliminată în timpul proteolizei limitate (partea întunecată a diagramei) și la două regiuni corespunzătoare ale subunității B. Resturile de aminoacizi, implicate în formarea legăturilor de hidrogen (distanța 2,3-3,4 Å), sunt arătate prin sfere.

Ar trebui stipulat, că termenul "Gln155" este utilizat, în mod convențional, în continuare pentru a desemna resturile de Gln, care coincid după poziția lor în alinierea multiplă vastă a  $\alpha$ -catenelor globulinelor 11S.



Fig. 3.14. Numărul legăturilor de hidrogen dintre subunitățile A şi B, formate cu participarea aminoacizilor subunităților A ale globulinelor 11S cu structura cuaternară cunoscută, în regiunea β-barrel-β-strand J (simbolurile albe) și dincolo (simbolurile negre).
Pe ordonată – numărul mediu de legături de hidrogen pe subunitate. Numerotarea resturilor de aminoacizi pe abscisă corespunde globulinei 11S 1fxz din soia în alinierea multiplă ale secvențelor de aminoacizi ale globulinelor 11S 1fxz, 2d5h, 3fz3, 3kgl şi 3ksc (Tabelul 3.2).



Fig. 3.15. Legăturile hidrogen (linii punctate) între restul de aminoacid Gln155 al subunității A și resturile Asn389, Gln426 și Thr446 ale subunității B în structura cuaternară a globulinei 11S 2d5h din soia.



Fig. 3.16. Restul de aminoacid Gln155 din regiunea dintre β-strendurile *I* și *J* ale subunității A unit prin legături de hidrogen cu resturile Asn389, Gln426 și Thr446 din structura subunității B a globulinei 11S 2d5h din soia.

Analiza secvențelor de aminoacizi ale globulinelor 11S, cu ajutorul programului BLAST utilizând în calitate de șablon secvența ginnacinei și globulinei 11S 1fxz din soia, arată un nivel ridicat de conservatism în regiunile structurilor primare ale acestor proteine, prezentat în Figura 3.17. Astfel, restul de aminoacid Gln155 dintre  $\beta$ -strendurile *I* și *J*, menționat mai sus (Figurile 3.14-3.16), este invariabil în toate globulinele 11S, de la *Ginkgo biloba* până la *Helianthus annuus* evolutiv avansat. Resturile Asn și Asn/His (dintre  $\beta$ -strendurile *C* și *D*), Gln (dintre  $\beta$ -strendurile *G* și *H*) și Thr ( $\beta$ -strendul *I*) din domeniul C-terminal al subunității B sunt practic invariabile (mai puțin de 3% substituții) (Figura 3.17). Prin urmare, se poate presupune că cel puțin în multe globuline 11S, restul invariabil Gln155 al subunității A joacă un rol important în formarea interacțiunilor intersubunitare, formând până la cinci legături de hidrogen cu trei resturi de aminoacizi dintr-o regiune vastă a  $\beta$ -barrelui ( $\beta$ -strendurile *CDEFGHI* ale subunității B). Eventual, apariția acestor legături este o etapă inițială necesară în formarea structurii trimerice a globulinelor 11S.

G	ln155																
	$\downarrow$	С	D	l	Ξ		I	=		G			Н			Ι	
2d5h	NQLD	NLNA <b>N</b> S	VIYVI	RGKGR	VRVV	NCQGN	AV <u>FD</u>	<u>GEL</u> R	RGQI	LVV	P <b>Q</b> NF	VVA	EQG	<u>G</u> EQ(	GLE	YV	/FK <b>T</b> H
1fxz	NQLD	NL <b>N</b> ANS	IIYAI	NGRAL	IQVV	NCNGE	RVFD	GELQ	EGRV	LIV	P <b>Q</b> NF	'VVA	ARS	QSDI	NFE	YVS	SFK <b>T</b> N
3ksc	NQLD	NLNA <b>N</b> S	IIYAI	KGRAR	LQVV	NCNGN	TVFD	GELE	AGRA	LTV	P <b>Q</b> NY	AVA	AKSI	LSDI	RFS	YVA	AFK <b>T</b> N
3fz3	NQLD	NV <b>N</b> A <b>H</b> S	VVYVI	RGNAR	VQVV	NENGD.	AILD	QEVQ	QGQI	FIV	P <b>Q</b> NH	IGVI	QQA	GNQ(	GFE	ΥFΖ	AFK <b>T</b> E
3kgl	NQLD	NANA <b>N</b> A	VLYVI	DGEAH	VQVV	NDNGD	RVFD	GQVS	QGQI	LSI	P <b>Q</b> GF	SVV	KRA'	TSE(	QFR	WIE	EFK <b>T</b> N
	****	* **::	::*.	*••	::**	* :*:	::*	::	*:	: :	**	*	:.	.:	:	:.	***•
	А							В									

**Fig. 3.17. Alinierea secvențelor de aminoacizi ale globulinelor 11S cu structura cuaternară cunoscută** (Tab. 3.2),înconjurată de restul de aminoacid conservativ Gln155 al subunității A, care formează legături de hidrogen cu resturile subunității B dintr-o regiune largă de structuri secundare (β-strendurile *C-I*). Structurile secundare ale globulinei 11S 2d5h din soia sunt subliniate.

Remarcăm, că pentru prima dată este descrisă specificitatea necondiționată a restului Gln155 din punctul de vedere al participării sale la stabilizarea interacțiunilor intersubunitare în structura cuaternară a globulinelor 11S.

#### 3.3. Concluzii la capitolul 3.

Rezultatele acestui studiu indică faptul că:

1. Ginnacina nativă este insensibilă la proteoliza "una câte una".

2. Scindarea profundă a ginnacinei prin mecanismul "una câte una" începe abia doar după atingerea unui anumit nivel al proteolizei limitate. Astfel, proteoliza limitată a ginnacinei are rolul de reglare, inițiind degradarea masivă a acestei proteine prin mecanismul "una câte una". Prin urmare rolul proteolizei limitate în inițierea degradării profunde a globulinelor de rezervă din semințe (*in vivo* și *in vitro*) a apărut la nivel evolutiv Ginkgophyta, reprezentanții străvechi ai spermatofitelor.

3. Capacitatea globulinelor 11S din semințe de a degrada masiv prin mecanismul "una câte una", dobândită numai după finalizarea proteolizei limitate relativ profunde, a apărut în stadiile incipiente ale evoluției acestor proteine (Figura 3.7), ale căror structuri primare și superioare sunt foarte conservative [2, 37]. Prin urmare, se poate presupune că rolul de reglare al proteolizei limitate în inițierea degradării profunde a globulinelor 11S s-a păstrat în timpul evoluției la proteinele din familia leguminelor.

# 4. PROTEOLIZA CA MODALITATE DE REDUCERE A ALERGENICITĂȚII GLOBULINELOR DE REZERVĂ DIN ARAHIDE

În prezent, până la 30% din populația lumii suferă de boli alergice, dintre care o parte importantă este alergia alimentară. Alergene sunt multe proteine de rezervă din semințe, în special globulinele de rezervă din arahide și din soia, utilizate pe scară largă în industria alimentară. După cum rezultă din analiza datelor din literatură (capitolul 1), modalitatea de reducere a nivelului de alergenicitate a globulinelor de rezervă din semințe poate fi proteoliza lor limitată.

#### 4.1. Materialul și metodele de cercetare

**Izolarea și purificarea globulinelor 7S și 11S din arahide.** Pentru izolarea alergenului 7S, Ara h1 și alergenului 11S, Ara h3 făina degresată din cotiledoanele de arahide (*Arachis hypogaea L.*) a fost extrasă cu soluția tampon A (0,05 M Tris-HCl, pH 8,0, NaN<sub>3</sub> 0,02%, EDTA 1 mM). Ara h1 și Ara h3 din extract au fost precipitate cu sulfat de amoniu. Precipitatele obșinute au fost dizolvate în 3,0 M NaCl în soluția tampon A și apoi soluțiile au fost introduse într-o coloană (2,6 x 9 cm) cu fenil-Sepharose CL-4B (Pharmacia Biotech, Suedia), echilibrată cu 3,0 M NaCl în soluția tampon A.

Proteoliza globulinelor 7S și 11S. Preparatele purificate de globulină 7S și 11S au fost hidrolizate cu papaină și tripsină (Sigma Life Science, SUA) la 30°C. Amestecul de reacție a conținut 5 mg/ml substrat și 50  $\mu$ g/ml enzimă (în cazul globulinei 7S) și fie 5  $\mu$ g/ml sau 10  $\mu$ g/ml enzimă (în cazul globulinei 11S) în soluția tampon B (0,05 M Tris-HCI pH 8,0, ajustat cu NaCl la puterea ionică 0,5, conținând 0,02% NaN<sub>3</sub>). Reacția a fost stopată prin adăugarea acidului tricloracetic (ATA) până la concentrația finală de 5%. În unele cazuri, reacția a fost stopată prin adăugarea E-64 (trans-epoxysuccinnyl-L-leucylamido-(4-guanidino)butan, Sigma, Life Science, SUA) până la concentrația finală de 10  $\mu$ M, și hidrolizatele au fost cercetate prin gelfiltrare pe Sephacryl S-300 High Resolution (Pharmacia Biotech, Suedia) în soluția tampon B (coloană de 0,9 x 75 cm).

**Electroforeza și analiza electroforegramelor.** Produsele macromoleculare ale proteolizei cu papaină și tripsină au fost cercetate prin SDS-PAGE (electroforeză în gel de poliacrilamidă a proteinei disociate cu dodecilsulfat de sodiu) în gel de 15%, în prezența și absența 2mercaptoetanolului (ME), în sistemul tampon Laemmli [106]. Au fost folosiți markerii moleculari PageRuler, Thermo Scientific (Lituania). Electroforegramele au fost scanate (ImageScanner III, GE Healthcare) și au fost analizate cu ajutorul programei Phoretix 1D Gel Analzsis v.5.10.

**Modelarea structurii cuaternare și metodele de analiză computerizată.** Modelarea omoloagă a structurii terțiare și cuaternare a globulinelor 7S și 11S din arahide a fost realizată cu ajutorul programei SWISS-Model [91, 92], folosind ca matriță structura cristalină a subunității sale pdb|3s7e și pdb pdb|3c3v, respectiv.

Progamul SDAP pentru a căuta și caracteriza alergenii 7S și 11S.

Programa DeepView/Swiss-PdbViewer v.3.7 a fost utilizată pentru construirea diagramelor panglică ale structurii terțiare ale Ara h1 și Ara h3, și pentru alinierea structurală a globulinelor 7S. Alinierea secvențelor complete ale globulinelor 7S a fost realizată folosind programa Clustal Omega.

Suprafața unui rest aminoacid X din structura proteinei accesibilă la solvent (ASA), exprimată în Å<sup>2</sup>, a fost determinată cu ajutorul programei <u>http://cib.cf.ocha.ac.jp/bitool/ASA/</u>. Valoarea relativă a ASA a fost exprimată în % din accesibilitatea restului X în tripeptida GXG. Analiza ASA a fost realizată utilizând structura model a trimerului Ara h1 și a hexamerului Ara h3.

Identificarea punctelor de scindare probabile ale Ara h1 în timpul proteolizei s-a efectuat așa cum s-a descris anterior, utilizând rezultatele electroforezei (compoziția fragmentelor polipeptidice și masele lor moleculare aparente conform SDS-PAGE, transformarea consecutivă a fragmentelor mari în fragmente mai mici, identificate prin modificarea intensității benzilor corespunzătoare), precum și următoarea informație de mai jos.

1. Rezultatele comparării structurilor terțiare ale Ara h1 și altor globuline 7S și ale Ara h3 și altor globuline 11S, a căror proteoliză limitată este descrisă [96], permit *a priori* determinarea poziției celor mai probabile regiuni de scindare ale Ara h1. Accesibilitatea sporită la solvent a acestor regiuni în structura Ara h1 și altor globuline 7S și în structura Ara h3 și altor globuline 11S poate fi considerată ca un indice al sensibilității lor potențiale la proteoliza limitată [2].

2. Punctele de clivare estimate în regiunile potențial sensibile ale moleculei Ara h1 și Ara h3 trebuie să corespundă specificității primare de substrat a papainei [105] (resturile Arg, Lys, Leu și Gly în poziția P1 în absența restului Val în poziția P1') și tripsinei [105] (resturile Arg, Lys în poziția P1 în absența restului Pro în poziția P1') Exista două variante ale subunităților Ara h1 din semințele de arahide (P43237 și P43238), ale căror moleculele mature aproape nu se deosebesc una de alta după masa moleculară (respectiv, 61,6 și 62,0 kDa) și similare după secvență (mai puțin de 3% din substituții neconservative și deleții). În această lucrare, pentru determinarea maselor moleculare ale fragmentelor am folosit secvența P43238, cea mai detaliat studiată variantă a Ara h1.

## 4.2. Izolarea și purificarea globulinei 7S.

În timpul salifierii globulina 7S, spre deosebire de globulina 11S, precipită la o concentrație relativ ridicată a sulfatului de amoniu (fracția 70-100% de saturație, Figura 4.1).



Fig. 4.1. SDS-PAGE extractului brut din semințele de arahide (pista 1) și fracțiile precipitate cu sulfat de amoniu în intervalul de concentrație (în % de saturație): 0-50% (pista 3), 50-70% (pista 5) și 70-100 % (pista 6). Pistele 2 și 4 arată compoziția proteinelor din supernatant la concentrația de sulfat de amoniu de 0-50% (2) și 50-70% (4) de saturație.

La cromatografia ulterioară pe coloana cu fenil-Sepharose CL-4B (Figura 4.2a), după îndepărtarea fracției neadsorbite în condițiile inițiale (3M NaCl), globulina 7S este eluată la concentrația NaCl 2 M (Figura 4.2b).



Fig. 4.2. Cromatografia pe coloana cu fenil-Sepharose CL-4B (a) și SDS-PAGE (b) a preparatului de globulină 7S din arahide, parțial purificat prin salifiere cu sulfat de amoniu. Linia orizontală indică regiunea de eluare a globulinei 7S. *1* – preparatul, parțial purificat prin salifiere; *2* și *3* - fracțiile eluate la cromatografia cu NaCl 1M și respectiv NaCl 2M.

# 4.3. Proteoliza limitată a Ara h1, globulina 7S din arahide

Structura globulinei 78 din arahide și regiunile secvenței sale de aminoacizi potențial sensibile la atacul papainei. Cu câteva excepții, cele două variante ale structurii subunității Ara h1, globulina 78 din arahide (pdb|3s7e [30] și pdb|3smh [31]), sunt tipice pentru globulinele de rezervă 78 din semințe, de tipul convicilinelor, care conțin în molecula matură prelungirea N-terminală nestructurată cu o sensibilitate crescută la atacul proteolitic [102]. În subunitatea  $\alpha'$  a globulinei 78 din soia, similară după structura primară cu Ara h1, prelungirea N-terminală este cea mai sensibilă la proteoliza limitată cu tripsină și este complet eliminată deja în stadiile inițiale ale reacției după scindarea legăturilor peptidice cu participarea restului de arginină din regiunea C-terminală a acestei prelungiri [15]. În aceeași regiune structurală a Ara h1 restul Arg87 formează o legătură peptidică cu restul Asn88 (Figura 4.3), care are o accesibilitate sporită la solvent (97%). În conformitate cu aceste observații, este probabil, că acțiunea inițială a papainei asupra Ara h1 duce la scindarea legăturii Arg87-Asn88, eliminând complet prelungirea N-terminală.



# Fig. 4.3. Diagramele panglică ale structurii terțiare ale domeniilor N-terminal (a) și Cterminal (b) în subunitățile globulinei 78, Ara h1, din arahide [113]

Liniile punctate arată regiunile nestructurate ale subunităților, potențial sensibile la proteoliza limitată: *a* - extensia N-terminală; *b* - bucla între α-helixul *h3* și β-strendul *J*' al domeniului Nterminal; *c* - linkerul interdomeniu; *d* - bucla între β-strendurile *E* și *F* a domeniului C-terminal; *e* - extensia C-terminală. Resturile Lys și Arg în poziția P1 a legăturii peptidice, care probabil sunt scindate de papaină în subunitățile Ara h1, sunt notate cu săgeți. Liniile groase arată poziția epitopilor IgE 4-9 (regiunea a) și 13 (regiunea b), care sunt scindate la hidroliza cu papaină

Una din caracteristicile specifice ale Ara h1 este prezența în molecula sa nu doar a prelungirii N-terminale, ci și a prelungirii C-terminale nestructurate relativ extinse (regiunea *e*, Figura 4.3), în porțiunea N-terminală a căreia este prezent restul Arg502 cu o accesibilitate sporită la solvent. Prin urmare, legătura peptidică Arg502-Pro503 este potențial sensibilă la atacul papainei. Scindarea estimată a legăturilor Arg87-Asn88 și Arg502-Pro503, care elimină atât prelungirea Nterminală, cât și C-terminală, trebuie să fie însoțită de formarea fragmentului cu o masă moleculară de 47,2 kDa (fragmentul *1*, Tabelul 4.1).

**Proteoliza globulinei 7S din arahide cu papaină.** În timpul hidrolizei cu papaină cel mai mare dintre fragmentele Ara h1 (fragmentul *I* cu masa moleculară aparentă de 48,7 kDa, Figura 4.4), care predomină cantitativ în etapele inițiale ale reacției, cu o probabilitate ridicată corespunde secvenței Asn88-Arg502 (Tabelul 4.1) [114].

NT		Masa moleculară, kDa					
Nr.	Regiunea de secvența	Prezisă	SDS-PAGE				
1	<i>a-e</i> N88-r502	47.17	48.7*				
2	a-d r54-r395	39.06	39.6				
3	<i>a-d</i> N88-r395	35.08	35.5				
4	<i>a-c</i> N88-k302	24.70	25.6				
5	<i>c-e</i> g303-r502	22.48	24.5*				
6	<i>a-b</i> N88-r272	21.40	21.0				
7	<i>a-b</i> R97-r264	19.25	19.4				
8	<i>b-d</i> s273-r395	13.70	14.0				
9	<i>d-e</i> r396-r502	12.10	12.8*				
10	<i>d-e</i> e412-r502	10.14	11.7*				
9a	<i>c-d</i> g303-r411	12.36	12.8				

Tabelul 4.1 Legături peptidice scindate de papaină din subunitățile Ara h1.

Notă: Literele mici arată resturile de aminoacizi, care aparțin regiunii nestructurate a cel puțin uneia dintre subunități. Numerotarea resturilor de aminoacizi corespunde moleculei mature a Ara h1 [16]. Asterixurile indica masele moleculare aparente, ceva mai mari datorită prezenței în fragment a restului de Asn437 glicozilat.

Orice alte căi alternative de formare a fragmentului *l* cu masa moleculară de aproximativ 49 kDa sunt practic imposibile, deoarece sunt asociate cu scindarea legăturilor peptidice din regiunile structural organizate inaccesibile ale moleculei. De menționat că masa moleculară aparentă a fragmentului *l* este puțin exagerată din cauza că restul de Asn437 din regiunea C- terminală a Ara h1 este glicozilat. De exemplu, în produsele de proteoliză a subunității  $\beta$  a globulinei 7S din soia masa moleculară aparentă a fragmentelor, care conțin un rest de asparagină glicozilată, este mărită cu 1,9-2,0 kDa [14].

Originea majorității celorlalte fragmente Ara h1 (Figura 4.4) poate fi descrisă cu o primă aproximare (Figura 4.5), reieșind din masele moleculare aparente obținute conform electroforezei (Tabelul 4.1) și din ipoteza formării lor în rezultatul scindării în regiunile nestructurate a, b, d și e(Figura 4.3), regiuni potențial sensibile la proteoliză. Fragmentul 2 este format în paralel cu fragmentul 1 în rezultatul clivării în regiunea d și îndepărtării inițiale incomplete a regiunii Nterminale a prelungirii a. Fragmentul 3 este format din fragmentul 1 (clivarea aceleiași regiuni d) și din fragmentul 2 după înlăturarea restului prelungirii N-terminale a. Fragmentele 6 și 8 sunt formate din fragmentul 3 în rezultatul scindării în regiunea b. Fragmentul 7 este format din fragmentul 6 în legătură cu scurtarea N-terminală și C-terminală. În final, fragmentul 10 este format din fragmentul 9 datorită scurtării N-terminale.



**Fig. 4.4. SDS-PAGE a alergenenului 7S, Ara h1, în timpul hidrolizei cu papaină** Linia de jos arată timpul de reacție (h). Pe partea stângă sunt indicate masele moleculare ale markerilor. Pe dreapta este arătată numerotarea fragmentelor.



Fig. 4.5. Schema ipotetică de formare a fragmentelor Ara h1 în timpul hidrolizei cu papaină

Linia de sus arată structurile secundare în secvențele subunităților.

Calea ipotetică de formare a fragmentelor de mai sus este descrisă în concordanță cu prezența în sectoarele presupuse de scindare a resturilor de aminoacizi cu o accesibilitatea sporită la solvent (Figura 4.3), care corespund specificității de substrat a papainei [105].

Lungimea regiunii *b*, potențial sensibilă la proteoliză, în domenul N-terminal a Ara h1, reprezintă încă una dintre particularitățile sale individuale, care și determină modul de formare a fragmentelor polipeptidice *6*, *7* și *8* descrise mai sus. Astfel, în subunitățile globulinei 7S din soia regiunea corespunzătoare este relativ scurtă și insensibilă la acțiunea papainei și tripsinei. Prin urmare, proteoliza în partea centrală a subunităților  $\alpha'$  și  $\beta$  ale globulinei 7S din soia se limitează la scindarea linkerului interdomenic [14, 15] (corespunzând regiunii *c* din subunitățile Ara h1).

În structura Ara h1 în regiunea c sunt prezente resturi de aminoacizi cu o accesibilitate sporită la solvent (Figura 4.3). Prin urmare, este posibil că fragmentele 4 și 5 nemenționate anterior sunt formate din fragmentul 1, după clivărea legăturii Lys302-Gly303 din partea centrală a regiunii c. Evident că fragmentele 4 și 5 (ca și fragmentul 1) sunt intermediare, deoarece secvențele lor conțin regiunile sensibile b și d, respectiv. Clivajul în aceste zone poate duce la apariția unor cantități suplimentare ale fragmentelor finale 7 și 10 și la apariția fragmentului 9a (Tabelul 4.1 și Figura 4.5), similar după mobilitate cu fragmentul 9.

Se pare că există două căi independente de proteoliză a Ara h1, datorită clivajului preferențial în etapele inițiale ale procesului de una dintre cele două regiuni nestructurate (regiunea

b sau regiunea c, Figura 4.3 și 4.5) potențial sensibile la atacul proteolitic. Ar trebui menționat faptul că în oligomerul pdb|3smh accesibilitatea la solvent a acestor zone într-o varietate de subunități nu este aceeași. În plus, regiunea c poate fi considerată nestructurată condițional, deoarece oricare dintre resturile de aminoacizi ale acestei regiuni sunt depistate în componența structurii cristaline a unei sau altei subunități pdb|3smh. Astfel, disponibilitatea potențială la atacul proteolitic a regiunilor b și c în diferite subunități ale oligomerului Ara h1 poate fi inegală, ceea ce determină un mod particular de proteoliză a acestora.

**Posibilități de reducere a alergenicității Ara h1 prin proteoliză limitată.** Pe baza scenariului ipotetic de mai sus (Figura 4.5) și poziției epitopilor IgE identificați în subunitățile Ara h1 [77], poate fi evaluat potențialul de a reduce alergenicitătea acestei proteine prin proteoliza sa limitată cu papaină (Figura 4.6).



Fig. 4.6. Secvența de aminoacizi a globulinei 7S, Ara h1, din arahide Literele mici prezintă regiunile nestructurate în structura cristalină a Ara h1 (pdb|3smh). Resturile de aminoacizi din regiunile nestructurate, care corespund specificității de substat a papainei [105] sunt subliniate. În fiecare rînd cifrele de jos corespund poziției epitopilor IgE 4-21, identificați în molecula matură Ara h1 [77]; cifrele de sus corespund numărului restului de aminoacid în poziția P1 a legăturii peptidice scindate de papaină. În subunitatea globulinei 7S din arahide Ara h1, care se caracterizează prin cea mai mare alergenicitate dintre globulinele de rezervă din semințe, s-a relevat prezența a 21 de epitopi IgE [77]. Epitopii IgE 1–3 aparțin secvenței N-terminale a subunității Ara h1, eliminate în timpul procesingului proteolitic [16]. A treia parte a celor 18 epitopi rămași în molecula matură a Ara h1 (epitopii IgE 4-9, Figura 4.7) aparțin regiunii N-terminale nestructurate a și sunt eliminați la acțiunea inițială a papainei. Această zonă este bogată în resturi de arginină și aparent este distrusă rapid cu formarea peptidelor scurte, ceea ce ar trebui să reducă alergenicitatea Ara h1. Este probabilă, de asemenea, distrugerea în timpul proteolizei limitate cu papaina a regiunii nestructurate b, care conține epitopul IgE 13.

Subunitățile globulinei 7S din arahide în compoziția moleculei sale hetero-trimerice sunt extrem de conservative. Este foarte probabil ca epitopii IgE, identificați în Ara h1, să fie de asemenea prezenți în alte subunități ale acestei proteine: indicii PD corespunzători nu depășesc 2,5 [115]. Astfel, proteoliza limitată cu papaină duce la o scădere parțială a alergenicității întregii molecule hetero-oligomere a globulinei 7S din arahide, datorită îndepărtării a mai mult de o treime din epitopii IgE.

Trebuie remarcat faptul, că accesibilitatea relativ sporită la solvent a regiunii  $\alpha$ -helixurilor *h1-h3*, care este de 3,5 ori mai mare decât ASA restului domeniului N-terminal al Ara h1. Prin urmare, este tentant să încercăm să găsim condiții pentru proteoliza limitată, care distrug această regiune potențial sensibilă a  $\alpha$ -helixurilor, unde sunt prezenți epitopii IgE 11 și 12 (Figura 4.6).

**Epitopi IgE potențiali în secvențele alergenilor 7S**. Probabilitatea prezenței epitopilor IgE, similari celor identificați în Ara h1, în secvențele alergenilor 7S din alte plante poate fi apreciat cantitativ cu ajutorul indicelui PD ((Property-Based Peptide Similarity Index for Two Sequences) [119]. Odată cu creșterea diferențelor între porțiunile secvențelor comparate ale Ara h1 și a altor proteine, valoarea indicelui PD crește de la zero (secvențele sunt identice) pînă la valoarea maximă egală cu 10, mai sus de care prezența epitopilor IgE corespunzători în proteina cercetată este puțin probabilă [119].

Conform acestei evaluări, epitopii 4-9 (prelungirea variabilă N-terminală) și 13-14 (Figura 4.3 și 4.6) sunt specifici numai pentru Ara h1. Secvențele tuturor celorlalți epitopi ai Ara h1 sunt omoloage și structural echivalente cu epitopii potențiali (PD de la 1,91 pînă la 6,07), descoperiți în alți alergeni 7S. În plus, epitopii 16-20 ai Ara h1 coincid la alinierea structurală cu o porțiune largă a domeniului C-terminal al alergenului 7S din linte Len c 1.0101, unde au fost identificate patru secvențe suprapuse ale epitopilor IgE [78].

La utilzarea în calitate de matrice a fiecăruia dintre epitopii Ara h1, cele mai mici valori ale indicelui PD au fost obținute preponderent pentru alergenii 7S din patru specii de plante - soia, linte, lupin și mazăre furajeră. Același grup de plante cu cea mai mare alergenicitate potențială este detectată prin compararea valorilor medii ale indicelor PD, obținute pentru alergeni 7S din nouă specii de plante (Figura 4.7. A). În cele din urmă, același grup de alergeni 7S formează un cluster comun cu Ara h1, la o analiză filogenetică a secvențelor complete ale domeniilor alergenilor 7S. (Figura 4.7). Astfel, imunoreactivitatea încrucișată între alergenii 7S din arahide și alte plante (în special, soia), apropiați după structura primară, este foarte probabilă.



Fig. 4.7. Valorile medii ale indicelui PD obținute prin compararea secvențelor regiunilor omoloage ale epitopilor 10-12 și 15-21 ai Ara h1 și alergenilor 7S ai altor plante (A), și analiza filogenetică a secvențelor domeniilor (418 poziții de aliniere) alergenilor 7S (B). Numerele de sus ale ramurilor dendrogramei corespund suportului statistic al clusterelor (% din 1000 replicații).

### 4.4. Izolarea și purificarea globulinei 11S.

La SDS-PAGE, în prezența 2-mercaptoetanolului, proteinelor extractului brut din semințele de arahide, se observă cinci benzi principale ale globulinelor de rezervă: banda globulinelor 7S, grupul din trei benzi ale  $\alpha$ -catenelor globulinei 11S și banda  $\beta$ -catenelor globulinei 11S (Figura 4.8). Partea principală a globulinei 11S se precipită în intervalul de

concentrație a sulfatului de amoniu 30-40% de saturație (pista 4). Această fracție a fost utilizată pentru purificarea ulterioară a globulinei 11S.

Cea mai mare parte a impurităților de globulină 7S se precipită la o concentrație de sulfat de amoniu peste 50% de saturație. La SDS-PAGE, în absența 2-mercaptoetanolului, în extractul brut și în fracțiile sale  $\alpha$ - și  $\beta$ -catenele globulinei 11S, unite printr-o legătură disulfurică formează benzi similare după mobilitate cu banda globulinei 7S (Figura 4.7, pistele 6-9).



Fig. 4.7. SDS - PAGE a extractului brut din semințele de arahide (pistele 1 și 6) și fracțiile precipitate cu sulfat de amoniu în intervalul de concentrație (în % de saturație)
0-20% (pistele 2 și 6), 20-30% (pistele 3 și 7), 30-40% (pistele 4 și 8) și 40-50% (pistele 5 și 9) în prezența (pistele 1-5) și în absența (pistele 6-9) 2-ME.

La cromatografia hidrofobă pe fenil-Sepharose CL-4B a preparatului de globulină 11S, parțial purificat prin salifiere, după îndepărtarea fracției care nu a fost adsorbită în condițiile inițiale (NaCl 3,0M), eluția ulterioară s-a efectuat prin micșorarea consecutivă a concentrației de NaCl (Figurile 4.8 și 4.9). Fracția 4, eluată la concentrația de NaCl 0,5 M, a fost utilizată pentru studiul ulterior.



Fig. 4.8. Cromatografia pe fenil-Sepharose CL-4B a preparatului de globulină 11S, parțial purificat prin salifiere cu sulfatului de amoniu (fracția 30-40% de saturație) 1-4 – fracțiile cercetate prin electroforeză. Fracția 4 a fost utilizată pentru studiul ulterior.



Fig. 4.9. SDS-PAGE a preparatului de globulină 11S, parțial purificat prin salifiere cu sulfatul de amoniu (pistele 1 și 6) și fracțiile eluate prin cromatografie pe fenil-Sepharose cu NaCl 1M (fracția 1-pistele 2 și 6, fracția 2-pistele 3 și 7, fracția 3-pistele 4 și 8) și NaCl 0,5M (fracția 4-pistele 5 și 9) în prezența (pistele 1-5) și în absența (pistele 6-9) 2-ME.

#### 4.5. Proteoliza limitată a Ara h3, globulină 118 din arahide.

Structurile primară și superioare ale globulinei 11S din arahide, Ara h3. Alergenul Ara h3 din arahide aparține familiei globulinelor 11S din semințe, relativ conservativă [38]. Structura hexamerică a globulinelor 11S este formată printr-o combinație aleatorie a mai multor variante de subunități (nu mai puțin de opt, în cazul globulinei 11S din arahide [116]). Fiecare dintre subunități este sintetizată într-un singur lanț polipeptidic, scindat posttranslational în  $\alpha$ - și  $\beta$ -catene, care formează respectiv domeniile N- și C-terminal structural echivalente. Fiecare domeniu constă din  $\beta$ -barrel format din opt  $\beta$ -strenduri antiparalele de bază (*BCDEFGHI*) și o serie de  $\beta$ -strenduri suplimentare, legate cu o grupă de  $\alpha$ -helixuri *h1, h2* și *h3* (Figura 4.10 și 4.11). Domeniile C-terminale sunt conservative, în timp ce domeniile N-terminale conțin trei inserții hidrofile nestructurate: *l* - între strendurile *E* și *F*, *2* – între  $\beta$ -barrel și grupul de  $\alpha$ -helixuri și *3* linkerul interdomenic alungit (Figura 4.12) [116].

În Ara h3, legumina din arahide, au fost identificați patru epitopi IgE [80] (Figura 4.11): primul epitop în regiunea N-terminală a  $\beta$ -barrelui ( $\beta$ -strendul *A* și  $\alpha$ -helix *h0*), al doilea și al treilea epitop, respectiv în regiunea  $\alpha$ -helixului *h1/h2* și în regiunea adiacentă  $\beta$ -strendului *J'*, și al patrulea epitop în regiunea hidrofilă C-terminală. De remarcat că trei dintre cei patru epitopi IgE din Ara h3 sunt localizați în regiunea C-terminală a  $\alpha$ -catenei, în afara  $\beta$ -barrel-ui.

IS <u>FRQ</u> QPEENACQ	)FQRLNAQRPE	NRIESEG	GY <b>ietwn</b>	PNNQEFECAG	/ALSRLVLRH	NALRRPFY
Z		A'	A	h0	В	С
SNA <u>PQEIFIQQG</u>	RGYFGLIFPG <b>C</b> E	PST <u>YEEP</u> .	<u>AQ</u> Qgrry ′	qsqrpprrlqe <b>1</b>	edqsq <u>QQQI</u> ت <b>ت</b>	OSHQKV <u>HRF</u>
NEGDLIAVPTGV2	AFWLYNDHDTD <b>G</b>	VVAVSLT H	<u>DTNNN</u> DN <b>I</b>	QLDQFPRR <u>FNI</u> J	LAGNHEQE <u>FI</u> h1	LRYQQqsrq L'
srrrslpyspysp	oqsqprqeere <b>2</b>	fsprgqh	srrerag	qeeeheg <b>GNII</b> h	SGFT <u>PEFL</u> 1 h2	AQAFQVDDR
QIVQNLRGENESE h3	EEQG <u>AIVTV</u> RG J'	GLRILSP	<b>drkr</b> gad	eeeey <b>dedeye</b>	eydeedrrro 3	<b>y</b> rgsrgsgn

Fig. 4.10. Secvența de aminoacizi și structurile secundare ale  $\alpha$ -catenei globulinei 118 din arahide (pdb|3c3v), formând  $\beta$ -barrel-ul, din  $\beta$ -strendurile de bază antiparalele *BCDE* și *FGHI*,  $\beta$ -strendurile suplimentare (*A'*, *A*, *E'*, *F'*, *J*, *J'* și *Z*) și  $\alpha$ -helixurile (de bază *h1*, *h2*, *h3* și suplimentare *h0* și *h1'*). Literele mici indică regiunile nestructurate *1*, *2* și *3* în cristalele 3c3v. Secvențele epitopilor IgE a Ara h3 și restul Cys88 între  $\beta$ -strendurile *E* și *E'*, implicat în formarea legături disulfidice între catenele  $\alpha$ - și  $\beta$  sunt indicate cu bold.

În molecula oligomeră a leguminelor subunitățile sunt unite în trimeri; trimerii a două subunități se combină într-un hexamer al proteinei mature. În structurile trimerice și hexamerice inserțiile hidrofile *1-3* (Figura 4.12) formează la suprafața moleculei straturi hidrofile [96] cu o accesibilitate ridicată la solvent (Figura 4.13) și, prin urmare, potențial sensibile la atacul proteolitic.



Fig. 4.11. Diagrama panglică a structurii terțiare a α-catenei (A) și a β-catenei (B) leguminei 3c3v *Arachis hypogaea* [117]

Baza structurii α- și β-catenelor leguminei din arahide și a altor legumine îl constituie βbarrel cu β-strenduri *BCDEFGHI*, suplimentat cu β-strendurile *A'*, *A* și *J*, *J'* și α-helixul *h0*. βstrendul *Z* este specific α-catenei leguminei [37]. α-helixul *h1'* suplimentar este prezent doar în unele din leguminele cu structura terțiară cunoscută [38]. Liniile punctate arată poziția inserțiilor hidrofile *1*, *2* și *3* în secvența α-catenei.



**Fig. 4.12. Diagrama panglică a structurii cuaternare a leguminei 3c3v din arahide** *a* – hexamerul subunităților A-F; *b* – trimerul subunităților A-C. Zonele întunecate pe diagrame descriu structura model a inserțiilor hidrofile *1-3* în una din α-catene.



Fig. 4.13. Resturile de aminoacizi cu accesibilitate sporită la solvent (ASA> 90%) în structura hexamerică model a globulinei 11S din arahide, Ara h3, sunt prezente exclusiv în α-catene și aproape exclusiv în straturile hidrofile de suprafață 1, 2 și 3, potențial sensibile la proteoliză. Simbolurile negre arată resturile, ce corespund specificității de substrat a papainei [105]. Săgeata indică poziția restului Cys88 implicat în formarea legăturii disulfidice

între  $\alpha$ - și  $\beta$ -catene.

**Proteoliza globulinei 11S din arahide cu papaină.** SDS-PAGE, în prezența ME, a preparatului purificat al globulinei 11S din arahide (Figura 4.14, A) detectează trei benzi majore și o bandă minoră cu masele moleculare aparente apropiate de cele calculate pentru cele patru variante principale ale α-catenelor acestei proteine (Tabelul 4.2.). β-catenele diferitelor subunități ale globulinei 11S din arahide diferă puțin una de alta după masa moleculară (20,69 ± 0,4 kDa) și formează doar o singură bandă electroforetică [116].

Numarul da acces	or an entre itataa	α-catene						
Numarur de acces	αp-subumatea	secventa	media	SDS-PAGE				
AAM46958 AAR02860	59.696 59.457	38.835 38.837	38.84	39.0				
AAD47382 ABF93402 3c3v AAG01363 ABL14270	58.849 58.574 58.485 58.223 58.214	37.929 37.921 37.963 37.638 37.570	37.80	37.8				
ADQ53859	56.284	35.614	35.61	35.6				
AAU21492	52.513	31.973	31.97	32.3				

Tabelul 4.2. Masele moleculare (kDa) a celor patru tipuri de subunitățiale globulinei 11S din arahide

În globulinele 11S din semințe  $\alpha$ - și  $\beta$ -catenele sunt prezente în cantități echimolare; calculate pentru Ara h3, cantitățile molare relative ale sumei celor patru tipuri de  $\alpha$ -catene (11,3% + 17,3% + 19,3% + 2,0% = 49,9%) și  $\beta$ -catenelor (50,1%), în limita erorii, sunt egale. Prin urmare, se poate presupune că benzile minore x și y (Figura 4.14, A) corespund impurităților din preparatul purificat al globulinei 11S, dar nu produselor hidrolizei parțiale a catenelor ei în timpul purificării.

În timpul proteolizei globulinei 11S cu papaină,  $\beta$ -catenele rămân intacte, iar conținutul  $\alpha$ catenelor se micșorează și apar fragmentele lor (Figura 4.14). Compararea rezultatelor SDS-PAGE ale produselor proteolizei în prezența și absența 2-ME (+ME și –ME, respectiv, Figura 4.15 a) permite separarea fragmentele  $\alpha$ -catenelor în două grupe. Fragmentele grupului K (K1-K3) sunt legate de  $\beta$ -catene prin legătură disulfidică și, prin urmare, nu pot fi detectate în absența ME. În schimb, comportamentul fragmentelor grupei N (N1-N4) la SDS-PAGE nu depinde de prezența sau absența ME. Menționăm că fragmentul N1 nu este detectat la SDS-PAGE în prezența ME (Figura 4.14 a), deoarece el coincide după mobilitate cu  $\beta$ -catenele intacte (Figura 4.15 a).



Fig. 4.14. SDS-PAGE globulinei 11S din arahide intacte si produsele de hidroliza cu papaina, in prezenta ME

Rîndul de jos arată timpul de reacție (min). În partea dreaptă sunt indicate masele moleculare ale markerilor (kDa). A și B - hidroliza în raportul enzimă/substrat de 1:1000 și 1:500, respectiv.



#### Fig. 4.15. Proteoliza globulinei 11S din arahide cu papaina

a - SDS-PAGE în prezenta (+ME) și absenta (-ME) 2-ME (20 min. de reacție, enzimă/substrat 1:1000). Fragmentele N1 și N4, care coincid după mobilitate cu β-catenele și fragmentul K3 respectiv, se găsesc numai în absența 2-ME. b - Schema reacțiilor succesive ale proteolizei α-catenei. 1, 2 și 3 - sectoarele potențial sensibile la proteoliză cu accesibilitate sporită la solvent (Fig. 4.14). Săgeata indică poziția restului de Cys 88 implicat în formarea legăturii disulfidice între α- și β-catene.

Consecutivitatea probabilă a evenimentelor proteolizei globulinei 11S cu papaină este prezentată în Figura 4.15 b. Reacția începe cu scurtarea C-terminală a  $\alpha$ -catenei, care este identificată prin apariția fragmentelor K1 și K2 (scindarea în regiunile sensibile *3* și, respectiv, *2*). Imediat după aceasta, scindarea buclei între strendurile *E'* și *F'* (regiunea sensibilă *1*) duce la formarea fragmentului K3 și a unei serii de fragmente N: N1 din fragmentul K1 și N2-N4 din fragmentul K2. Conform acestui scenariu al proteolizei, fragmentele K1, K2 și N1, N2 sunt intermediare; întradevăr, ele dispar la o proteoliză mai profundă (Figura 4.15 b). Fragmentele finale N3 și N4, probabil, corespund câtorva fragmente omoloage ale  $\alpha$ -catenelor diferitelor subunități ale globulinei 11S, ce diferă după masa moleculară. Sensibilitatea la atacul proteolitic, cel puțin a regiunii C-terminale a  $\alpha$ -catenelor, este caracteristică pentru toate globulinele 11S din semințe studiate în această privință [2].

În conformitate cu scenariul descris mai sus (Figura 4.15 b), proteoliza limitată a globulinei 11S din arahide cu papaină scindeaza regiunea C-terminală a  $\alpha$ -catenei, unde sunt prezenți trei dintre cei patru epitopi IgE (Figura 4.10). Astfel, gradul de alergenicitate al globulinei 11S poate fi redus substanțial prin proteoliza sa limitată cu papaină.

**Proteoliza globulinei 11S din arahide cu tripsină.** Compoziția fragmentelor, formate în timpul hidrolizei globulinei 11S din arahide cu papaina specific relativ scăzut, este suficient de complexă, ceea ce face dificilă stabilirea legăturilor peptidice scindabile. Informațiile corespunzătoare au fost obținute la studierea hidrolizei globulinei 11S din arahide cu tripsina îngust specifică, descrisă mai jos.

Hidroliza globulinei 11S cu tripsină are loc suficient de rapid, și după 5 minute la raportul enzimă/substrat 1:1000 se observă formarea produsului, care în continuare rămâne practic neschimbat (Figura 4.16 și 4.17), constituit din  $\beta$ -catenele intacte și trei fragmente ale  $\alpha$ -catenelor cu masele moleculare aparente de 13,3 kDa (fragment N1) și 12,7 kDa (fragmente N2 și K2). Compararea rezultatelor SDS-PAGE ale produselor proteolizei în prezența și absența ME (+ME și –ME, respectiv, Figura 4.16) indică existența a două tipuri de fragmente ale  $\alpha$ -catenelor. Comportamentul fragmentelor N1 și N2 nu depinde de prezența sau absența ME; ele sunt reținute în molecula de proteină parțial hidrolizată numai datorită interacțiunilor necovalente. Fragmentul K2, legat printr-o legătură disulfidică cu  $\beta$ -catena, și în absența ME formează banda K2 $\beta$  (Figura 4.16).



Fig. 4.16. SDS-PAGE globulinei 11S din arahide intacte si produselor proteolizei cu tripsină, in prezenta (+ME) si absenta (-ME) ME

Rândul de jos arată timpul de reacție (min.). În partea dreaptă sunt indicate masele moleculare ale proteinelor markeri. Benzile x și y corespund impurităților în preparatul inițial al globulinei 11S (vezi mai sus).

După cum s-a menționat mai sus (Tabelul 4.2.), în componența moleculei heterohexamerice a globulinei 11S din arahide intră patru variante principale de subunități, similare după structura primară. Evident, fiecare dintre fragmentele N1, N2 și K2 este eterogen; fiecare dintre ele aparține regiunilor omoloage, dar neidentice, ale secvențelor diferitelor subunități.

În timpul hidrolizei cu tripsină, conținutul relativ al β-catenelor intacte, exprimat în procente molare, este constant (Figura 4.17, graficul 1) și reprezintă o treime (32,9 ± 0,9%) din suma a trei componente, detectate la SDS-PAGE în prezența ME (Figura 4.16): (fragmentul N1 (Figura 4.17, graficul 2), fragmentele K2/N2 (Figura 4.17, graficul 3) și β-catenele (Figura 4.17, graficul 1). Evident, că fragmentele K2 după conținutul relativ (în procente molare) nu diferă de β-catenele, cu care sunt legate printr-o legătură disulfidică. De aceea, conținutul relativ a fragmentelor N2 (în procente molare) este egal cu diferența dintre conținutul sumar al fragmentelor K2 și N2 (Figura 4.17, graficul 3) care formează o bandă comună (Figura 4.16, +ME) și β-catene (Figura 4.17, graficul 1). Conform acestei scheme, produsul proteolizei limitate a globulinei 11S cu tripsină constă din cantitățile echimolare a trei componente: β-catenele N1 și N2 se formează în timpul proteolizei a două tipuri de subunități ale globulinei 11S din arahide, ce diferă doar puțin după structura primară [118].
Masa moleculară medie a produsului cu masa moleculară mare a tripsinolizei globulinei 11S, calculată ca suma maselor  $\beta$ -catenelor, fragmentului K2 și masa moleculară medie a fragmentelor N1 și N2, este constantă și alcătuiește  $80,9\% \pm 0,1\%$  (Figura 4.17, graficul 5) din masa moleculară inițială a proteinei (ultima este calculată ca suma maselor  $\beta$ -catenelor și masa moleculară medie a tuturor celor opt variante principale ale  $\alpha$ -catenelor [116]).



Fig. 4.17. Analiza cantitativă a proteolizei globulinei 11S din arahide cu tripsină Abscisa - timpul de reacție (min.). Ordonata – conținutul, în procente molare, al produselor polipeptidice ale reacției (1-4) și masa moleculară a subunităților, în procente din cea inițială (5, 6). *1* – β-catenele intacte; *2* - fragmentul N1; *3* - suma fragmentelor N2 și K2; *4* - fragmentul N2; *5* și *6* – masa moleculară a subunităților, calculată după conținutul relativ, în procente molare, al masei moleculare a polipeptidelor proteinei inițiale și amestecului de reacție (*5*) și după modificarea conținutului relativ al β-catenelor în timpul reacției (*6*).

În timpul proteolizei limitate a globulinei 11S cu tripsină  $\beta$ -catenele rămân intacte. Acest lucru dă posibilitatea de a determina modificarea masei moleculare relative a subunităților în timpul reacției printr-o metodă alternativă independentă [9]. În acest caz, M<sup>t</sup>/m<sup>0</sup> = P<sub>β</sub><sup>0</sup>/P<sub>β</sub><sup>t</sup>, unde m<sup>0</sup> și M<sup>t</sup> - masa moleculară inițială și după timpul de reacție t, respectiv; P<sub>β</sub><sup>0</sup> și P<sub>β</sub><sup>t</sup> - conținutul de masă relativ al β-catenelor în amestecul de reacție inițial și după timpul de reacție t, respectiv. Astfel calculată masa moleculară relativă a proteinei reziduale în limita erorii este constantă (81.8 ± 1.8%) (Figura 4.17, graficul 6) și este apropiată de valoarea obținută prin metoda anterioară (80.9 ± 0.1%). Aceasta confirmă compoziția polipeptidică a subunităților globulinei 11S din arahide după finisarea proteolizei limitate cu tripsină, descrisă mai sus [118].

Rezultatele de mai sus, împreună cu informația descrisă anterior privind prezența în structura moleculei Ara h3 a regiunilor potențial sensibile la atacul proteolitic (Figura 4.13), permit

identificarea celor mai probabile puncte de clivaj ale  $\alpha$ -catenelor globulinei 11S din arahide (Figura 4.16). Jumătatea N-terminală a  $\beta$ -barrelui, care conține cisteina, implicată în formarea legăturii disulfidice între  $\alpha$ - și  $\beta$ -catene, conservativă în structura primară a subunităților globulinei 11S din arahide și alte globuline 11S, nu conține resturi de aminoacizi cu o accesibilitate sporită la solvent (Figura 4.9) și, prin urmare, nu este sensibilă la proteoliza limitată. Evident, fragmentul K2, legat covalent cu  $\beta$ -catenele, aparține jumătății N-terminale a  $\beta$ -barrelui și se formează în rezultatul scindării în regiunea inserției hidrofile nestructurate *1* (Figura 4.13), potențial sensibilă la atacul proteolitic în structura Ara h3 și a altor globuline 11S [2]. E destul de probabil că fragmentul K2 se formează prin scindarea legăturii peptidice r109-r110 (Figura 4.18), deoarece masa sa moleculară aparentă, conform datelor SDS-PAGE (12,7 kDa), coincide cu masa care corespunde regiunii N-terminale a tuturor celor opt variante principale (Tabelul 4.2.) de subunități ale globuline 11S din arahide (12,69 ± 0,06 kDa).



### Fig. 4.18. Schema secvențelor α-catenelor în subunitățile 3c3v și ABL14270 ale globulinei 118 din arahide

Sunt arătate structurile secundare în  $\alpha$ -catena subunității pdb|3c3v și în  $\alpha$ -catenele ambelor subunități – secvențele complete de aminoacizi (cu litere mici), potențial sensibile la proteoliză, ale inserțiilor hidrofile *I*(intre  $\beta$ -strendurile E' și F'), *2* (intre  $\alpha$ -helixurile *h1*' și *h1*) și *3* (regiunea C-terminală). Asterixul marchează poziția restului de cisteină implicat în formarea legăturii disulfidice între  $\alpha$ - și  $\beta$ -catene. Cu bold sunt evidențiate: structurile secundare care limitează sectoarele nestructurate *I* și *2*; resturile de aminoacizi care corespund specificității de substrat a tripsinei, pentru care este arătată o accesibilitate sporită la solvent (> 80%) în structura model a hexamerului 3c3v [117]. Săgețile corespund celor mai probabile puncte de scindare ale  $\alpha$ catenelor. Italic sunt arătate masele moleculare ale fragmentelor (kDa), calculate după segmentul corespunzător al secvenței (numerotarea resturilor de aminoacizi corespunde subunităților mature ale moleculei). În numeroase studii [2] se arată că proteoliza globulinei 11S din semințele germinate și *in vitro* începe cu scindarea C-terminală a inserției hidrofile nestructurate 3 (Figura 4.13). E destul de probabil ca globulina 11S din arahide nu este diferită după acest caracter de celelalte globuline 11S: în structura model a hexamerului Ara h3 [117], unicul rest de aminoacid cu o accesibilitate relativă la solvent mai mare de 100%, este Arg282, ce aparține  $\beta$ -strendului J' (Figura 4.18). Menționăm că regiunea corespunzătoare este aproape identică (doar tolerează substituțiile conservative Arg/Lys) în secvențele tuturor subunităților globulinei 11S din arahide [117].

Fenomenul formării fragmentelor N1 și N2, care se deosebesc după masa moleculară a produselor finale ale proteolizei limitate a subunităților înalt conservative ale globulinei 11S din arahide, poate fi ilustrat prin compararea secvențelor  $\alpha$ -catenelor ale subunităților 3c3v și ABL14270 (Figura 4.18).

După îndepărtarea regiunii C-terminală a sectorului *3* al subunităților 3c3v și ABL14270 și scindarea legăturii r109-r110, în secvențele reziduale ale acestor subunități rămâne numai inserția hidrofilă extinsă *2* potențial sensibilă la proteoliză (Figura 4.18). În subunitatea 3c3v în partea centrală a inserției *2*, este prezent restul r213 cu o accesibilitate sporită la solvent (> 80%). Scindarea legăturii corespunzătoare (r213-q214) este posibilă, deoarece masa moleculară a segmentului r110-r213 al secvenței 3c3v (12,32 kDa) și a fragmentului N2 (12,7 kDa conform datelor SDS-PAGE) sunt apropiate.

Structurile primare (evident, și cele secundare) ale subunităților 3c3v și ABL14270 diferă puțin una de alta. Cu toate acestea, în subunitatea ABL14270 din regiunea centrală a inserției 2, restul Arg (restul r213 în subunitatea 3c3v) este înlocuit cu restul Gly (Figura 4.18). În acest sens, punctul de scindare al secvenței ABL14270, cel mai probabil, este deplasat, ceea ce duce la formarea fragmentului N1 (r110-r225), cu masa moleculară de 13,61 kDa (13,3 kDa conform datelor SDS-PAGE).

Conform scenariului descris, produsul final al proteolizei limitate a globulinei 11S din arahide duce la distrugerea rapidă (până la formarea produselor care nu sunt detectabile prin electroforeză) a întregii regiuni C-terminale a  $\alpha$ -catenelor, inclusiv a  $\alpha$ -helixurilor *h1*, *h2* și *h3*. Această regiune cuprinde trei dintre cei patru epitopi IgE identificați în secvența de aminoacizi ale Ara h3 [80] (Figura 4.10). Astfel, proteoliza limitată cu tripsină duce la o reducere substanțială a nivelului de alergenicitate a globulinei 11S din arahide.

### 4.6. Imunoreactivitatea încrucișată a globulinelor de rezervă

În secvențele de aminoacizi ale globulinelor de rezervă 7S și 11S din semințele de arahide [77, 80] și de soia [17, 81] au fost identificați determinanții antigenici (epitopii IgE), care determină imunoreactivitatea lor crescută. O parte din acești epitopi IgE aparțin regiunilor de secvență, în cadrul fiecăreia dintre cele două familii extinse de globulinele 7S și 11S, care formează structuri secundare echivalente [2, 7]. Structura primară a acestor regiuni este foarte conservativă, ceea ce indică prezența probabilă a epitopilor IgE, similari celor identificați în globulinele de rezervă din arahide și soia, în regiunile omoloage secvențelor globulinelor 7S și 11S din semințele altor plante.

Secvențele de aminoacizi ale globulinelor 7S și 11S, care provin dintr-un precursor evolutiv comun, sunt omoloage [29, 43]. Aceasta indică o posibilă imunoreactivitate încrucișată nu doar a globulinelor 7S și globulinelor 11S în cadrul fiecăreia dintre aceste familii, dar și între aceste familii.

Programul SDAP oferă posibilitatea de a evalua cantitativ probabilitatea prezenței epitopilor IgE într-o anumită proteină [119]. Metoda se bazează pe compararea caracteristicilor (a peste 200) fizico-chimice ale fiecăreia dintre secvențele de aminoacizi ale epitopilor IgE, identificați în proteina alergenă, cu fiecare dintre aminoacizii apropiați în structura primară a regiunii secvenței proteinei, care este studiată. Gradul de similaritate a secvenței comparate este estimat cu ajutorul indicelui PD (Property-Based Peptide Similarity Index for Two Sequences). Pe măsură ce diferențele dintre secvențele comparate cresc, indicele PD crește de la zero (secvențele sunt identice) până la o valoare limită, egală cu zece, deasupra căreia prezența epitopului IgE corespunzător în proteina cercetată este puțin probabilă [119].

**Strategia de cercetare.** Pentru a evalua imunoreactivitatea încrucișată potențială a globulinelor de rezervă, se recomandă de a utiliza în calitate de șablon numai secvențele epitopilor IgE, identificați în globulinele 7S și 11S, care prezintă imunoreactivitate încrucișată reală, adică coincid parțial după poziție în structurile lor primare și terțiare. Această condiție este satisfăcută de epitopii IgE ai globulinelor 11S din arahide, Ara h3 și din soia Gly m G1, precum și ai globulinei 7S din arahide, Ara h1 (Figura 4.19).

Pentru a detecta imunoreactivitatea încrucișată a globulinelor de rezervă, a fost inițial efectuată o BLAST căutare, utilizând în calitate de șablon cele trei secvențe de aminoacizi cu o lungime de 21 resturi, prezentate în Figura 4.19. Cel puțin două dintre secvențele detectate (primul

nivel al BLAST căutărilor) sunt utilizate în calitatea de șabloane de nivel secundar pentru BLAST căutările ulterioare repetate.

			<u>h1</u> <u>h2</u>
Ara	h3	238	EG <b>GNI<u>F</u>SG<u>F</u>TPE<u>FL</u>AQ</b> AFQVD
Gly	G1	197	E <b>GGSILSGFTLEFLEHAFSV</b> D
Ara	h1	138	DQSS <u>YLQGF</u> SRNTLEAAFNAE
			* * * * * * * * * *

### Fig. 4.19. Imunoreactivitatea încrucișată a epitopilor IgE ai globulinelor 11S din arahide, Ara h3 (epitopul 2) și din soia Gly m G1 (epitopul 1), precum și ai globulinei 7S din arahide, Ara h1 (epitopul 11)

Secvențele epitopilor IgE sunt evidențiate cu bold. Resturile de aminoacizi, înlocuirea cărora duce la scăderea nivelului de alergenicitate a Ara h3 și Ara h1 [77, 80] sunt subliniate. În structurile cristaline ale tuturor celor trei proteine (Ara h3, pdb|3c3v; Gly m G1, pdb|1fxz; Ara h1, pdb|3smh), α-helixurile *h1* și *h2* după poziție și extensie coincid complet.

Pentru fiecare dintre cele trei colecții obținute, indicii PD au fost determinați utilizând în calitate de şablon secvențele a trei epitopi IgE (Figura 4.19). În concordanță cu rezultatele obținute, au fost întocmite trei colecții finale cele mai informative ale secvențelor cu valori relativ scăzute ale indicilor PD. În continuare, indicii PD au fost determinați pentru toate secvențele acestor trei colecții, folosind toate cele trei variante de epitopi IgE (Figura 4.19) ca șabloane în toate cele nouă combinații de perechi posibile [120].

**Rezultatele studiului** imunoreactivității potențiale a globulinelor de rezervă din arahide, soia și alte plante, prezentate în Figura 4.20, sunt documentate prin datele cantitative prezentate în Tabelul 4.3.

Probabilitatea prezenței unui epitop IgE similar cu cel identificat în Ara h3 (Figura 4.19) în globulinele 11S 1-9 este destul de mare: indicii PD corespunzători sunt mai mici de 4 (Figura 4.20, Tabelul 4.3). Este mult mai puțin probabilă (Tabelul 4.3) prezența în aceleași globuline 11S a epitopului IgE similar cu epitopul IgE al Gly m G1 (Figura 4.19, 4.20). Se observă aceeași legitate atunci cînd se analizează un set mai mare de secvențe ale globulinelor 11S (datele nu sunt prezentate). Diferențele în regiunile omoloage ale secvențelor epitopilor IgE ai Gly m G1 și ai altor globuline 11S studiate sunt evidente (Tabelul 4.3).

Probabilitatea prezenței unui epitop IgE, similar cu cel identificat în Ara h1 (Figura 4.19) cel puțin în cinci globuline 7S ale grupului 7S *a* (17-21, Figura 4.20), este suficient de mare: indicii PD corespunzători sunt mai mici de 3 (Tabelul 4.3). Este evidentă absența completă a

imunoreactivității încrucișate a epitopului IgE al Ara h1 (Figura 4.19) și a porțiunii omoloage a secvențelor globulinelor 11S 1-9 (Figura 4.20, Tabelul 4.3).



Fig. 4.20. Imunoreactivitatea încrucișată a globulinelor de rezervă din semințe în regiunile omoloage ale structurilor primare din regiunea α-helixurilor h1 și h2 ale domeniilor Nterminale (Fig. 4.21)

Numerele din linia de jos corespund globulinelor 11S 1-9 și globulinelor 7S ale grupelor 7S *b* (10-16) și 7S *a* (17-21), prezentate în Tabelul 4.3. Ca șablon au fost utilizați epitopii IgE identificați în globulinele 11S din soia Gly m G1 și din arahide Ara h3, precum și în globulina 7S din arahideAra h1 (Fig. 4.19).

La întocmirea, pa baza BLAST căutărilor, a unei colecției de globuline, înrudite cu secvența epitopului IgE din Gly m G1 (Figura 4.19), se detectează nu doar globuline 11S, ci și globuline 7S din grupul 7S *b* (Figura 4.20, Tabelul 4.3). Indicii PD corespunzători sunt relativ mari - de la 5 până la 7 (Figura 4.20, Tabelul 4.3). Cu toate acestea, trebuie de avut în vedere faptul că Gly m G1 și Ara h1 aparțin la două familii diferite de globuline de rezervă. De aceea, este foarte probabil că imunoreactivitatea încrucișată presupusă între globulinele 7S (grupul 7S b) și globulinele 11S 1-9, prezentată în Figura 4.20, va fi confirmată la identificarea ulterioară a epitopilor IgE ai unui număr mai mare de globuline de rezervă în regiunea  $\alpha$ -helixurilor *h1* și *h2* ale domeniilor N-terminale.

Dacă această ipoteză este adevărată, datele prezentate în Figura 4.20 indică prezența probabilă în secvențele globulinelor 7S a două tipuri de epitopi IgE în regiunea  $\alpha$ -helixurilor h1 și h2: epitopii 7S b și, respectiv, 7S a, care detectează și nu detectează imunoreactivitate încrucișată cu globulinele 11S.

7S din arahide, Ara h1 (Figura 1). Resturile de aminoacizi din globulinele 7S și 11S studiate, identice cu cele prezente în fiecare dintre cei trei epitopi, sunt evidențiate cu bold. Resturile, înlocuirea cărora duce la o scădere a alergenicității Ara h3 și Ara h1, sunt subliniate.

	<b>2</b>		Ara h3 Ara h3	PD	Gly Gl	- Da	Ara hl	DD
	Numarul de accesar	e Specia	GNIESGETPEFLAQA	0.00	GGSILSGFTLEFLEHAFSV	0.00	SYLQGESRNT	0.00
				11S				
<u> </u>	1 kyp70740	Cajanus	GNIFSGFTPEFLeQA	1.06	<b>GGnIfSGFT</b> p <b>EFLE</b> q <b>A</b> ln <b>V</b>	3.59	nifs <b>GF</b> tpef	10.91
	2 otg11724	Helianthus	<b>GNIF</b> nGFTPEliAQS	2.68	a <b>G</b> n <b>I</b> fn <b>GFT</b> p <b>E</b> liags <b>F</b> n <b>V</b>	7.02	nifn <b>GF</b> tpel	10.37
	3 kfk42731	Arabis	<b>NNIFSGFaPEVLAQA</b>	2.69	qnn <b>I</b> f <b>SGF</b> ap <b>E</b> v <b>L</b> aq <b>AF</b> ki	7.95	nifs <b>GF</b> apev	11.35
	4 afq32288	Camelina	n <b>niftgfafeilaga</b>	2.85	qnn <b>I</b> ft <b>GF</b> ap <b>E</b> i <b>L</b> aq <b>AF</b> ki	8.08	nift <b>GF</b> apei	11.52
	5 aaa32778	Arabidopsis	n <b>nifngfapeilaga</b>	2.93	qnn <b>I</b> fn <b>GF</b> ap <b>E</b> i <b>L</b> aq <b>AF</b> ki	8.14	nifn <b>GF</b> apei	11.33
	6 eoa36574	Capsella	nNIFnGFaPE1LAQA	2.93	qnn <b>I</b> fn <b>GE</b> ap <b>E</b> i <b>L</b> aq <b>AF</b> ki	8.14	nifn <b>GE</b> apei	11.33
	7 xp_018449576	Raphanus	nNIFnGFaPEiLAQA	2.93	qnn <b>I</b> fn <b>GF</b> ap <b>E</b> i <b>L</b> aq <b>AF</b> ki	8.14	nifn <b>GF</b> apei	11.33
	8 esq36546	Eutrema	nNIFnGFaPEiLAkA	3.51	qnn <b>I</b> fn <b>GF</b> ap <b>E</b> i <b>L</b> ak <b>AF</b> ki	8.54	nifn <b>GF</b> apei	11.33
<u> </u>	9 cdy18193	Brassica	n <b>NIF</b> n <b>GF</b> a <b>P</b> Qi <b>LAQA</b>	3.63	qnn <b>I</b> fn <b>GF</b> apqi <b>L</b> aq <b>AF</b> ki	8.69	nifn <b>GF</b> apqi	10.73
				7S a				
	10 xp_015874370	Ziziphus	ts <b>IlSGF</b> sthi <b>L</b> en <b>A</b>	6.66	ptSILSGFsthiLEnAFnV	4.99	SiLsGFSthi	7.48
1	11 snq45153	Cannabis	as <b>IlSGF</b> ds <b>EiL</b> en <b>A</b>	6.43	pa <b>silsgf</b> ds <b>filfnAf</b> n <b>V</b>	5.19	<b>SilsGF</b> dsei	9.12
_	12 exb98567	Morus	tsv1 <b>SGF</b> dsEiLenA	6.56	ptSvLSGFdsEiLEnAFnV	5.29	<b>S</b> v <b>L</b> s <b>GF</b> dsei	9.22
_	13 oiw15764	Lupinus	as <b>IlSGF</b> s <b>PEiI</b> et <b>A</b>	7.46	pasillsGFspEilEtAFnV	5.30	SiLsGFSpei	8.43
	14 abk80758	Ficus	vs <b>iisgf</b> ds <b>f</b> ilen <b>A</b>	6.81	pv <b>siisgf</b> ds <b>eilenAf</b> n <b>v</b>	6.00	SiisGFdsei	9.89
	15 gau15615	Trifolium	vsvl <b>sGF</b> q <b>P</b> qi <b>L</b> es <b>A</b>	7.04	pvSvLSGFqpqiLEsAFnV	6.52	<b>s</b> v <b>l</b> s <b>gf</b> qpqi	9.17
	16 xp_011457328	Fragaria	ksvl <b>SGF</b> dsEiLtnA	7.12	pk SvLSGFdsEiLtnAFnV	6.85	<b>S</b> v <b>L</b> s <b>GF</b> dsei	9.22
				4 SL				
	17 aeb33713	Lupinus	sy <b>F</b> n <b>GF</b> srnt	10.76	syfn <b>GF</b> srnt	10.28	SYfnGFSRNT	1.82
-	18 bab64304	Glycine	sylq <b>GF</b> srni	10.99	sy <b>L</b> q <b>GF</b> srni	10.34	SYLQGFSRN	1.91
1	19 2cv6	Vigna	sylq <b>GF</b> skni	10.93	sy <b>L</b> q <b>GF</b> skni	10.00	SYLQGFSKNİ	2.63
(1	20 kyp48930	Cajanus	sylq <b>GF</b> skni	10.93	sy <b>L</b> q <b>GF</b> skni	9.68	SYLQGFSKNİ	2.63
CN	21 car78998	Lotus	syln <b>GF</b> srni	10.93	sy <b>L</b> n <b>GF</b> srni	9.43	SYLNGFSRNi	2.77
Notă: Ca	şablon au fost uzilizat	e secvențele ep	itopilor IgE ai globulin	elor 115	s din soia, Gly m G1 și din a	arahide,	Ara h3 și ai glo	bulinei 7

Tabelul 4.3. Evaluarea cantitativă a imunoreactivității încrucișate globulinelor de rezervă 7S și 11S în regiunile omoloage ale structurilor loi Dintre rezultatele comparării perechilor de secvențe ale celor trei epitopi IgE selectați din arahide și din soia cu regiunile omoloage ale secvențelor altor globuline de rezervă, prezentate în Tabelul 4.3, atrag atenția următoarele:

Toate secvențele, fără excepție, conțin o pereche de resturi superconservative Gly și Phe.
 Înlocuirea ambelor resturi în epitopul IgE din Ara h1 [78] și a restului Phe în epitopul IgE din Ara
 h3 [82] duce la scăderea nivelului de alergenicitate al acestor globuline.

2. Restul Leu, a cărui înlocuire în epitopul IgE din Ara h3 duce la scăderea alergenicității acestei proteine, este prezent în toate globulinele 7S din grupul 7S *b*. Acesta este un argument suplimentar în favoarea imunoreactivității încrucișate presupuse a globulinelor 11S și globulinelor 7S din grupul 7S *b*.

3. Cele două resturi Gly N-terminale ale epitopuluui IgE din Gly m G1 nu sunt caracteristice pentru secvențele globulinelor 7S din grupul 7S *b*. Dacă în calitate de șablon se utilizează secvența epitopului IgE din Gly m G1 scurtat cu două resturi N-terminale, atunci valoarea medie a indicelui PD pentru globulinaele 7S din grupul 7S *b* scade de la  $5,73\pm0,67$  până la  $4,36\pm0,50$ . Această observație reprezintă un alt argument în favoarea imunoreactivității încrucișate presupuse a globulinei 11S Gly m G1 și a globulinelor 7S din grupul 7S *b*.

Rezultatele de mai sus ale studiului privind imunoreactivitatea potențială a globulinei 11S Gly m G1 și a globulinelor 7S din grupul 7S b pot stimula experimente ulterioare pentru identificarea epitopilor IgE în globulinele 7S ale acestui grup.

### 4.7. Concluzii la capitolul 4.

1. A fost elaborată metoda optimă de izolare și purificare a globulinei 11S alergene, Ara h3, și a globulinei 7S alergene, Ara h1, din semințele de arahide, care include numai trei etape – extragerea, fracționarea cu SA și cromatografia hidrofobă. Prin urmare, preparatul proteic pur poate fi obținut cu un număr minim de operații, care permit obținerea proteinei în stare mult mai nativă, ceea ce este important pentru studierea proteolizei sale ulterioare.

2. S-a arătat scindarea la proteoliza limitată cu papaină și tripsină a regiunii C-terminale a domeniului N-terminal al globulinei 11S din arahide, unde sunt prezenți trei pătrimi dintre epitopii IgE identificați în această proteină. Astfel, proteoliza limitată a leguminei din arahide duce la o reducere semnificativă a alergenicității sale.

3. S-a constatat, că proteoliza limitată cu papaină distruge complet regiunea N-terminală nestructurată în structura primară a globulinei 7S din arahide, unde sunt prezenți o treime dintre epitopii IgE, identificați în această proteină. Prin urmare, perspectivele de reducere, prin proteoliza limitată, a alergenicității vicilinei din arahide, de asemenea, sunt favorabile.

4. Rezultatele studiului imunoreactivității încrucișate ale globulinelor de rezervă din semințe indică prezența stabilită (globulinele 11S din arahide și soia), precum și prezența probabilă a epitopilor IgE în regiunea C-terminală omoloagă înalt conservativă a domeniilor N-terminale ale globulinelor 11S și 7S. După cum s-a arătat anterior în studiul prezent (capitolul 3), scindarea acestei regiuni are loc în timpul proteolizei limitate inițiale a multor proteine de rezervă din semințe din familia globulinelor 11S. Aceasta arată perspectivele de reducere parțială prin proteoliză limitată a alergenicității nu doar a globulinelor 11S și 7S din arahide și din soia, ci și a altor globuline de rezervă ale acestor familii.

### CONCLUZII GENERALE ȘI RECOMANDĂRI

### **CONCLUZII GENERALE**

1. A fost stabilită originea globulinelor de rezervă din semințe, leguminelor și vicilinelor, din oxalat-decarboxilazele bidomenice bacteriene (capitolul 2). Concluzia se bazează pe rezultatele analizei evolutive a secvențelor de aminoacizi din colecția vastă de globuline de rezervă din semințe și proteinelor înrudite structural din bacterii, eucariotele inferioare, din plantele cu spori și plantele cu semințe [96].

2. Precursorii evolutivi străvechi **direcți** ai globulinelor de rezervă din semințe sunt proteinele legumino-similare din alga verde harofită *Klebsormidium nitens* (capitolul 2). Concluzia se bazează pe analiza tuturor etapelor succesive ale evoluției globulinelor de rezervă din semințe, pornind de la oxalat-decarboxilazele bacteriene și proteinele înrudite din alga verde clorofită *Coccomyxa subellipsoidea*, precum și pe rezultatele modelării computerizate a structurilor terțiare și cuaternare ale proteinei din alga verde *Klebsormidium nitens* [101].

3. Structurile primare și terțiare ale leguminei reflectă trăsăturile străvechi ale globulinelor de rezervă din semințe (capitolul 2). Concluzia se bazează pe topologia arborelui evolutiv al globulinelor de rezervă din semințe, precum și pe rezultatele analizei cantitative ale comparării secvențelor de aminoacizi și structurilor cristaline ale leguminelor și vicilinelor și precursorului lor din cianobacterii [96].

4. Proteoliza limitată a leguminei din semințele *Ginkgo biloba* (ginnacina) inițiază degradarea masivă prin mecanismul "una câte una". Rolul de reglare corespunzător al proteolizei limitate este caracteristic multor globuline de rezervă din familia leguminelor (capitolul 3). Prin urmare, mecanismul de reglare a degradării globulinelor de rezervă în timpul germinării semințelor, format deja la nivelul Ginkgophyta, a fost moștenit în timpul evoluției de globulinele de rezervă din semințe. Concluziile se bazează pe rezultatele studierii cineticii proteolizei limitate și "una cîte una" a ginnacinei și a analizei tuturor structurilor cristaline cunoscute ale leguminelor, precum și pe rezultatele analizei legăturilor de hidrogen intersubunitare responsabile de formarea structurii cuaternare a leguminelor [110, 111].

5. Proteoliza limitată a leguminei și a vicilinei din arahide duce la o scădere semnificativă a nivelului de alergenicitate a acestor proteine de rezervă (capitolul 4). S-a stabilit că proteoliza limitată cu papaină scindează complet regiunea N-terminală nestructurată în structura primară a vicilinei din arahide, unde sunt prezenți o treime din epitopii IgE identificați în această proteină. S-a arătat scindarea în timpul proteolizei limitate cu papaină și tripsină, a regiunii C-terminale a domeniului N-terminal al leguminei din arahide, unde sunt prezenți trei din cei patru epitopi IgE, identificați în această proteină [113-118, 120].

# RECOMANDĂRI

1. Proteoliza limitată poate fi recomandată ca o modalitate de a dezvolta metode eficiente de reducere semnificativă a nivelului de alergenicitate al globulinelor de rezervă din arahide și din soia, semințele cărora sunt utilizate pe scară largă în produsele alimentare [115].

# BIBLIOGRAFIE

- DUNWELL, J.M. Structure, function, and evolution of vicilin and legumin seed storage proteins. In: *Biotechnology of Biopolymers - From Synthesis to Patents* (Steinbüchel, A., Doi, Y., eds.), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005, p. 967-997. ISBN 978-3527311101.
- SHUTOV, A.D and WILSON, K.A. Seed storage globulins: their descent from bacterial ancestors and mechanisms of degradation. In: *Globulins: Biochemistry, Production and Role in Immunity* (Milford, S.D., ed.), Nova Science Publishers, New York, 2014, p. 71-104. ISBN 978-1-63117-781-1
- IVANCIUK, O., SCHEIN, C.H., BRAUN, W. SDAP: database and computational tools for allergenic proteins. In: *Nucleic Acid Res.*, 2003, vol.31, p. 359-362. ISSN 0305-1048. doi:10.1093/nar/gkg010
- 4. SHUTOV, A.D., BÄUMLEIN, H. Origin and evolution of seed storage globulins. In: *Seed Proteins* (P. R. Shewry & R. Casey, eds.), Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999, p. 543-561. ISBN 978-94-011-4431-5.
- DUNNWELL, J. M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. In: *Biotechnology Genet. Engeneering Rev.*, 1998, vol.15, p. 1-32. ISSN 0264-8725. <u>doi:org/10.1080/02648725.1998.10647950</u>
- SHUTOV, A.D., KAKHOVSKAYA, I.A. Evolution of seed storage globulins and cupin superfamily. In: *Mol. Biol.* (Moscow), 2011, vol.45, p. 579-585. ISSN 0026-8933.<u>doi:10.1134/S0026893311030162</u>
- 7. ADACHI, M. et al. Crystal structure of soybean proglycinin A1aB1b homotrimer. In: *J. Mol. Biol.*, 2001, vol.305, p. 291-305. ISSN 0022-2836. doi:10.1006/jmbi.2000.4310
- ADACHI, M. et al. Crystal structure of soybean 11S globulin: glycinin A3B4 homohexamer. In: *Proc. Natl. Acad. Sci. USA*, 2003, vol.100(12), p. 7395-7400. ISSN 0027-8424. doi:10.1073/pnas.0832158100
- SHUTOV, A. et al. Limited proteolysis regulates massive degradation of glycinin, storage 11S globulin from soybean seeds: An *in vitro* model. In: *Journal of Plant Physiology*, 2012, vol.169, p. 1227-1233. ISSN 0176-1617. <u>doi:10.1016/j.jplph.2012.06.004</u>
- RUDAKOVA, A.S., RUDAKOV, S.V., KAKHOVSKAYA, I.A., SHUTOV, A.D. 11S storage globulin from pumpkin seeds: regularities of proteolysis by papain. In: *Biochemistry* (*Moscow*), 2014, vol.79, p. 820-825. ISSN 0006-2979. doi:10.1134/S0006297914080100.
- ZAKHAROV, A. et al. Comparative study of the role of the major proteinases of germinated common bean (*Phaseolus vulgaris* L.) and soybean (*Glycine max* (L.) Merrill) seeds in the degradation of their storage proteins. In: *J. Exp. Bot.*, 2004, vol.55, p. 2241-2249. ISSN 0022-0957. <u>doi:10.1093/jxb/erh247.</u>
- 12. SENYUK, V. et al. Does an asparaginyl-specific cysteine endopeptidase trigger phaseolin degradation in cotyledons of kidney bean? In: *Eur. J. Biochem.*, 1998, vol.258, p. 546-558. ISSN 0014-2956.
- 13. JIVOTOVSKAYA, A.V. et al. Proteolysis of phaseolin in relation to its structure. In: *J. Agric. Food Chem.*, 1996, vol.44, p. 3768-3772. ISSN 0021-8561. doi:org/10.1021/jf9601291.
- SHUTOV, A.D. et al. Degradation of β-conglycinin β-homotrimer by papain: independent occurrence of limited and extensive proteolyses. In: *Biosci. Biotechnol. Biochem.*, 2013, vol.77, p. 2082-2086. ISSN 0916-8451. doi:10.1271/bbb.130440.

- SHUTOV, A. D. et al. Limited proteolysis of β-conglycinin and glycinin, the 7S and 11S storage globulins from soybean (*Glycine max* (L.) Merr.): structural and evolutionary implications. In: *Eur. J. Biochem.*, 1996, vol.241, p. 221-228. ISSN 0014-2956. doi:10.1111/j.1432-1033.1996.0221t.x.
- WICHERS, H.J., DE BEYER, T., SAVELKOUL, F.J., VAN AMERONGEN A. The major peanut allergen Ara h1 and its cleaved-off N-terminal peptide; possible implications for peanut allergen detection. In: *J. Agric. Food Chem.*, 2004, vol.52, p. 4903-4907. ISSN 0021-8561. doi:10.1021/jf0496970.
- 17. BEARDSLEE, T. A., ZEECE, M. G., SARATH, G., MARKWELL, J. P. Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3. In: *Int. Arch. Allergy Immunol.*, 2000, vol.123, p. 299-307. ISSN 1018-2438. doi:10.1159/000053642.
- 18. OSBORNE, T.B., CAMPBELL, G.F. The proteids of the pea, lentil, horse bean and vetch. In: *J. Amer. Chem. Soc.*, 1898, vol.20, p. 410-418.
- DERBYSHIRE, E., WRIGHT, D.J., BOULTER, D. Legumin and vicilin, storage proteins of legume seeds. In: *Phytochemistry*, 1976, vol.15, p. 3-24. ISSN 0031-9422. doi:org/10.1016/S0031-9422(00)89046-9.
- BOROTTO, K., DURE, L. The globulin storage proteins of flowering plants are derived from two ancestral genes. In: *Plant Mol. Biol.*, 1987, vol.8, p. 113-131. ISSN 0167-4412. doi:10.1007/BF00025323.
- CASEY, R. Distribution and some properties of seed globulins. In: *Seed Proteins* (P. R. Shewry & R. Casey, eds.), Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999, p. 159-169. ISBN 978-94-011-4431-5.
- 22. DANIELSSON, C.E. Seed globulins of the gramineae and leguminosae. In: *Biochem. J.*, 1949, vol.44, p. 387-400. ISSN 0264-602. <u>doi:10.1042/bj0440387.</u>
- 23. DANIELSSON, C.E. Investigation of vicilin and legumin. In: *Acta Chem. Scand.*, 1949, vol.3, p. 41-49. doi:10.3891/acta.chem.scand. 03-0041.
- 24. MARUYAMA, N. et al. Crystal structures of recombinant and native soybean β-conglycinin β homotrimers. In: *Eur. J. Biochem.*, 2001, vol.268, p. 3595-3604. ISSN 0014-2956. doi:org/10.1046/j.1432-1327.2001.02268.x.
- MARUYAMA, Y., MARUYAMA, N., MIKAMI, B., UTSUMIN, S. Structure of the core region of the soybean β-conglycinin α' subunit. In: J. Acta Cryst. Section D, 2004, vol.60(PT2), p. 289-297. ISSN 1399-0047. doi:org/10.1107/S0907444903027367.
- NIELSEN, N.C., NAM, Y.W. Soybean globulins. In: Seed Proteins (P. R. Shewry & R. Casey, eds.), Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999, p. 285-313. ISBN 978-94-011-4431-5.
- KOSHIYAMA, I., FUKUSIMA, D. Identification of the 7S globulin with β-conglycinin in soybean seeds. In: *Phytochenistry*, 1976, vol.15, p. 157-159. ISSN 0031-9422. doi:org/10.1016/S0031-9422(00)89075-5.
- KO, T.P. et al. Determination of three crystal structures of canavalin by molecular replacement. In: J. Acta Cryst. Section D, 1993, vol.49 (PT 5), p. 478-489. ISSN 1399-0047. doi:org/10.1107/S0907444993004056.
- 29. LAWRENCE, M.C. et al. Structure of phaseolin in 2.2 Å resolution. In: *J. Mol. Biol.*, 1994, vol.238, p. 748-776. ISSN 0022-2836. doi:10.1006/jmbi.1994.1333

- 30. CHRUSZEZ, M. et al. Structural and immunologic characterization of Ara h1, a major peanut allergen. In: J. Biol. Chem., 2011, vol.286(45), p. 39318-39327. ISSN 1083-351X. doi:10.1074/jbc.M111.270132.
- 31. CABANOS, C. et al. Crystal structure of the major peanut allergen Ara h 1. In: *Mol. Immunol.*, 2011, vol.49, p. 115-123. ISSN 0161-5890. <u>doi:10.1016/j.molimm.2011.08.004.</u>
- ITOH, T. et al. Structure of 8S alpha globulin, the major seed storage protein of mung bean. In: J. Acta Cryst. Section D, 2006, vol.62 (PT7), p. 824-832. ISSN 1399-0047. doi:org/10.1107/S090744490601804X.
- FUKUDA, T. et al. Characterization and crystallography of recombinant 7S globulins of Adzuki bean and structure-function relationships with 7S globulins of various crops. In: J. Agric. Food Chem., 2008, vol.56(11), p. 4145-4153. ISSN 0021-856. doi:10.1021/jf072667b.
- KO, T.P., NG, J.D., MCPHERSON, A. The three-dimentional structure of canavalin from jack bean (*Canavalia ensiformis*). In: *Plant Physiol.*, 1993, vol.101(3), p. 729-744. ISSN 1532-2548. doi:org/10.1104/pp.101.3.729.
- 35. JIN, T. et al. Crystal Structure of Korean Pine (*Pinus koraiensis*) 7S Seed Storage Protein with Copper Ligands. In: *J. Agric. Food Chem.*, 2014, vol.62(1), p. 222-228. ISSN 0021-8561. doi:10.1021/jf4039887.
- ADACHI, M. et al. Crystal structures and structural stabilities of the disulfide bond-deficient soybean proglycinin mutants C12G and C88S. In: J. Agric. Food Chem., 2003, vol.51(16), p. 4633-4639. ISSN 0021-8561. doi:10.1021/jf026065y.
- 37. JIN, T. et al. Crystal structure of Ara h3, a major allergen in peanut. In: *Mol. Immunol.*, 2009, vol.46(8-9), p. 1796-1804. ISSN 0161-5890. <u>doi:10.1016/j.molimm.2009.01.023</u>.
- TANDANG-SILVAS, M.R. et al. "Conservation and divergence on plant seed 11S globulins based on crystal structures. In: *Biochim. Biophys. Acta*, 2010, vol.1804(7), p. 1432-1442. ISSN 0304-4165. doi:10.1016/j.bbapap.2010.02.016.
- 39. JIN, T. et al. Crystal structure of prunin-1, a major component of the almond (Prunus dulcis) allergen amandin. In: *J. Agric. Food Chem.*, 2009, vol.57(18), p. 8643-8651. ISSN 0021-8561. doi:10.1021/jf9017355.
- TANDANG-SILVAS, M.R. et al. Expression, Purification And Preliminary Crystalliz Amaranth 11S Proglobulin Seed Storage Protein From Amaranthus Hypochondriacus L. In: J. Acta Cryst. Section F, 2010, vol.66(PT8), p. 919-922. ISSN 2053-230X. doi.org/10.1107/S1744309110021032.
- 41. ARGOS, P, NARAYAMA, SVL, NIELSEN, NC. Structural similarity between legumin and vicilin storage proteins from legumes. In: *EMBO J.*, 1985, vol.4(5)2, p. 1111-1117. ISSN 0261-4189.
- 42. GIBBS, P.E., STRONGIN, K.B., MCPHERSON, A. Evolution of legume storage proteins a domain common to legumins and vicilins is duplicated in vicilins. In: *Mol. Biol. Evol.* 1989, vol.6, p. 614-623. ISSN 0737-4038. <u>doi:10.1093/oxfordjournals.molbev.a040575.</u>
- SHUTOV, A. D. et al. Legumin-like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. In: *J. Mol. Evol.*, 1995, vol.41, p. 1057-1069. ISSN 0022-2844. doi:10.1007/bf00173187
- 44. BÄUMLEIN, H., BRAUN, H., KAKHOVSKAYA, I. A., SHUTOV, A.D. Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. In: *J. Mol. Evol.*, 1995, vol.41, p. 1070-1075. ISSN 0022-2844. doi:10.1007/bf00173188.

- 45. WOO, E. J. et al. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. In: *Nature Struct. Biol.*, 2000, vol.7, p. 1036-1040. ISSN 1476-4687. doi:10.1038/80954.
- 46. DUNWELL, J. M. & GANE, P. J. Microbial relatives of seed storage proteins; conservation of motifs in a functionally diverse superfamily of enzymes. In: *J. Mol. Evol.*, 1998, vol.46, p. 147-154. ISSN 0022-2844. doi:10.1007/pl00006289.
- DUNWELL, J. M., KHURI, S., GANE, P. J. Microbial relatives of the seed storage proteins of higher plants: construction of structure, and diversification of function during evolution of the cupin superfamily. In: *Microbiol Mol. Biol. Rev.*, 2000, vol.84, p. 153-179. ISSN 1092-2172. doi:10.1128/mmbr.64.1.153-179.2000.
- 48. KHURI, S., BAKKER, F.T., DUNWELL, J. M. Phylogeny, function and evolution of the cupins, a structurally conserved, functionally diverse superfamily of proteins. In: *Mol. Biol. Evol.*, 2001, vol.18, p. 593-605. ISSN 0737-4038. doi:10.1093/oxfordjournals.molbev.a003840.
- 49. DUNWELL, J. M. et al. Evolution of functional diversity in the cupin superfamily. In: *Trends Biochem. Sci.*, 2001, vol.26, p. 740-746. ISSN 0968-0004. doi:10.1016/s0968-0004(01)01981-8.
- 50. SHUTOV, A.D., VAINTRAUB, I.A. Degradation of storage proteins in germinating seeds. In: *Phytochemistry*, 1987, vol.26, p. 1557-1566. ISSN 0031-9422. doi:org/10.1016/S0031-9422(00)82245-1.
- 51. WILSON, K.A. Role of proteolytic enzymes in the mobilization of protein reserves in germinating dicot seeds. In: *Dalling MJ. Plant proteolytic enzymes*. Boca Raton, Florida: CRC Press Inc. 1986, vol. 2, p. 19-47.
- 52. MÜNTZ, K. Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. In: *J. Exp. Bot.*, 1996, vol.47, p. 605-622. ISSN 0022-0957. doi:org/10.1093/jxb/47.5.605.
- 53. ISHI, S. Legumain: asparaginyl endopeptidase. In: *Meth. Enzymol.*, 1994, vol.244, p. 604-615. ISSN 0076-6879. doi:org/10.1016/0076-6879(94)44044-1.
- MÜNTZ, K., BLATTNER, F.R., SHUTOV, A.D. Legumains a family of asparagine-specific cysteine endopeptigases involved in propolypeptide processing and protein breakdown in plants. In: J. Plant Physiol., 2002, vol.159, p. 1287-1293. ISSN 0176-1617. doi:org/10.1078/0176-1617-00853
- 55. MÜNTZ, K., SHUTOV, A.D. Legumains and their functions in plants. In: *Trends Plant Sci.*, 2002, vol.7, p. 340-344. ISSN 1360-1385.
- 56. HARA-NISHIMURA, I., INOUE, K., NISHIMURA, M. A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. In: *FEBS Lett.* 1991, vol.294, p. 89-93. ISSN 1873-3468. doi:10.1016/0014-5793(91)81349-d.
- HIRAWA, N., TAKEUCHI, Y., NISHIMURA, M., HARA-NISHIMURA, I. A vacuolar processing enzyme in maturing and germinating seeds: its distribution and associated changes during development. In: *Plant and Cell Physiology*, 1993, vol.34, p. 1197-1204. ISSN 0032-0781. doi:org/10.1093/oxfordjournals.pcp.a078541.
- 58. HARA-NISHIMURA, I., KINOSHITA, T., HIRAWA, N., NISHIMURA, M. Vacuolar processing enzymes in protein storage vacuoles and lytic vacuoles. In: *J. Plant Physiol.*, 2002, vol.152, p. 668-674. ISSN 0176-1617.

- 59. BELOZERSKY, M.A., DUNAEVSKY, Y.E., RUDENSKAYA, G.N., STEPANOV, V.M. Carboxyl proteinases from buckweat seeds. In: *Biochemistry (Moscow)*, 1984, vol.49, p. 401-407. ISSN 0006-2979.
- BELOZERSKY, M.A., SARBAKANOVA, S.T., DUNAEVSKY, Y.E. Aspartic proteinase from from wheat seeds: isolation, properties and action on glyadin. In: *Planta*, 1989, vol.177, p. 321-324. ISSN 0032-0935.
- 61. ABE, Y. et al. Asparaginyl endopeptidase of jack bean seeds. Purification, characterization and high utility in protein sequence analysis. In: *J. Biol. Chem.*, 1993, vol.268, p. 3525-3529. ISSN 1083-351X.
- 62. HARA, I., MATSUBARA, H. Pumpkin (*Cucurbita sp.*) seed globulin. VI. Proteolytic activities, appearing in germinating cotyledons. In: *Plant Cell Physiol.*, 1980, vol.21, p. 233-245. ISSN 0032-0781. doi.org/10.1093/oxfordjournals.pcp.a075997.
- 63. BOND, H.M., BOWLES, D.J. Characterization of soybean endopeptidase activity using exogenous and endogenous substrates. In: *Plant Physiol.*, 1983, vol.73, p. 345-350. ISSN 0176-1617. doi:10.1104/pp.72.2.345.
- 64. BELOZERSKY, M.A., DUNAEVSKY, Y.E., VOSKOBOYNICOVA, E. Isolation and properties of a metalloproteinase from buckweat (Fagopyrum esculentum) seeds. In: *Biochem. J.*, 1990, vol.272, p. 677-682. ISSN 0264-6021
- 65. SHUTOV, A.D. et al. Action of trypsin on soybean glycinin. Mixed-type proteolysis and its kinetics; molecular mass of glycinin-T. In: *Eur J Biochem.*, 1991, vol.199, p. 539-543. ISSN 0014-2956.
- 66. VAINTRAUB, I.A. Kinetics of co-operative proteolysis. In: *Nahrung*, 1998, vol.42, p. 59-60. ISSN 0027-769X.
- 67. RUPLEY, J. A. Susceptibility to attack by proteolytic enzymes. In: *Methods Enzymol.*, 1967, vol.11, p. 905-17. ISSN 0076-6879.
- 68. KAWAI, M. et al. Characterization of 30-kDa fragments derived from beta-conglycinin degradation process during germination and seedling growth of soybean. In: *Biosci. Biothech. Biochem.*, 1997, vol.61, p. 794-799. ISSN 0916-8451. doi: 10.1271/bbb.61.794.
- SEO, S., TAN-WILSON, A., WILSON, K.A. Protease C2, a cysteine endopeptidase involved in the continuing mobilization of soybean β-conglycinin seed proteins. In: *Biochim. Biophys. Acta*, 2001, vol.1545, p. 192-206. ISSN 0304-4165. <u>doi:org/10.1016/S0167-4838(00)00277-</u> <u>6.</u>
- 70. RUDAKOVA, A. et al. *In vivo* and *in vitro* limited proteolysis of phaseolin: facts, suggestions and problems. In: *Studia Universitatis*, seria *Ştiinţe ale naturii*, 2007, nr.1, p.101-111. ISSN 1814-3237.
- 71. ROTARI, V. I. et al. Proteinase A-like enzyme from germinated kidney bean seeds. Its action on phaseolin and vicilin. In: *Plant Physiol.*, 1997, vol.100, p. 171-177. ISSN 0032-0889.
- 72. WILSON, K.A., RIGHTMIRE, B.R., CHEN, J.C., TAN-WILSON, A.L. Differential proteolysis of glycinin and β-conglycinin polypeptides during soybean germination and seedling growth. In: *Plant Physiol.*, 1986, vol.82, p. 71-76. ISSN 0032-0889. <u>doi:10.1104/pp.82.1.71.</u>
- 73. MAKAEVA, E. et al. Sunflower seed 11S globulin: kinetics of papain limited and cooperative proteolysis. In: *Studia Universitatis*, seria *Ştiinţe ale naturii*, 2009, nr.1, p. 24-28. ISSN 1814-3237.

- 74. SHUTOV, A.D. et al. Limited proteolysis of oat 11S globulin by papain. In: *Studia Universitatis*, Seria *Ştiințe ale naturii*, 2014, nr.1, p. 85-90. ISSN 1814-3237.
- 75. RUDAKOVA, A. et al. Limited proteolysis controls massive degradation of glycinin, storage 11S globulin from soybean seeds. In: *Agrobiodiversitatea Vegetală în Republica Moldova: Evaluatea, Conservarea şi Utilizarea.* Materialele Simpozionului Național, Chişinău, 2008, p.396-402.
- 76. LAPTEVA, N. et al. Mixed-type papain proteolysis of cedar 11S globulin. In: *Studia Universitatis*, seria *Ştiinţe ale naturii*, 2010, nr.6, p. 9-13. ISSN 1814-3237.
- SHIN, D.S. et al. Biochemical and structural analysis of the IgE binding sites on Ara h 1, an abundant and highly allergenic peanut protein. In: J. Biol. Chem., 1998, vol.273, p. 13753-13759. ISSN 0021-9258. doi:10.1074/jbc.273.22.13753.
- VEREDA, A. et al. Identification of IgE sequential epitopes of lentil (Len c1) by means of peptide microarray immunoassay. In: *J. Allergy Clin. Immunol.*, 2010, vol.126, p. 596-601. ISSN 0091-6749. doi:10.1016/j.jaci.2010.06.023.
- 79. PELE, M. Peanut allergens. In: *Romanian Biothechnol. Lett.*, 2010, vol.15(2), p. 5204-5212. ISSN: 2248-3942.
- RABJOHN, P. et al. Molecular cloning and epitope analysis of the peanut allergen Ara h 3. In: J. Clin. Invest., 1999, vol.103, p. 535-542. ISSN 0021-9738. doi:10.1172/JCI5349.
- HELM, R. M. et al. A soybean G2 glycinin allergen. 1. Identification and characterization. In: Int. Arch. Allergy Immunol., 2000, vol.123, p. 205-212. ISSN 1018-2438. doi:10.1159/000024445
- KLEBER-JANKE, T. et al. Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. In: *Int. Arch. Allergy Immunol.*, 1999, vol.119, p. 265-274. ISSN 1018-2438. doi:10.1159/000024203.
- HELM, R. M. et al. A soybean G2 glycinin allergen. 2. Epitope mapping and threedimensional modeling. In. *Int. Arch. Allergy Immunol.*, 2000, vol.123, p. 213-219. ISSN 1018-2438. doi:10.1159/000024446.
- WILLISON, L.N. et al. Cloning, expression and patient IgE reactivity of recombinant Pru du 6, an 11S globulin from almond. In: *Int. Arch. Allergy Immunol.*, 2011, vol.156(3), p. 267-281. ISSN 1018-2438. doi:10.1159/000323887.
- BEYER, K. et al. Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. In: J. Allergy Clin. Immunol., 2002, vol.110(3), p. 517-523. ISSN 0091-6749. doi:10.1067/mai.2002.127434.
- PALOMARES, O. et al. Isolation and identification of an 11S globulin as a new major allergen in mustard seeds. In: *Ann. Allergy Astma Immunol.*, 2005, vol.94, p. 586-592. <u>ISSN</u> 1081-1206. <u>doi:10.1016/S1081-1206(10)61138-6.</u>
- 87. SHARMA, G.M. et al. Cloning and characterization of an 11S legumin, Car i4, a major allergen in pecan. In: J. Agric. Food Chem., 2011, vol.59, p. 9542-9552. ISSN 00218561. doi:org/10.1021/jf2017447.
- WANG, F. et al. Ana o2, a major cashew (*Anacardium occidentale* L.) nut allergen of the legumin family. In: *Int. Arch. Allergy Immunol.*, 2003, vol.132, p. 27-39. ISSN 1018-2438. doi:10.1159/000073262.

- 89. YOSHIOKA, H. et al. Expression and epitope analysis of the major allergenic protein Fag e1 from buckwheat. In: *Journal of Plant Physiology*, 2004, vol.161, p. 761-767. ISSN 0176-1617. doi:10.1016/j.jplph.2004.01.010.
- 90. VAN DE PEER, Y. & DE WACHTERR, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. In: *Comput. Appl. Biosci.*, 1994, vol.10, p. 569-570. ISSN 0266-7061. doi:10.1093/bioinformatics/10.5.569
- ARNOLD, K., BORDOLI, L., KOPP, J., SCHWEDW, T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. In: *Bioinformatics*, 2006, vol.22, p. 195-201. ISSN 1367-4803. <u>doi:10.1093/bioinformatics/bti770</u>
- 92. BIASINI, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. In: *Nucleic Acids Research*, 2014, vol.42, p. 252-258. ISSN 0305-1048. doi:10.1093/nar/gku340
- NIELSSON, S.S., LIENER, I.E. Degradation of the major storage protein of *Phaseolus vulgaris* during germination. Role of endogenous proteases and protease inhibitors. In: *Plant Physiol.*, 1984, vol.74, p. 494-498. ISSN 0032-0889. doi:10.1104/pp.74.3.494.
- 94. BASHA, S.M., BEEVERS, L. The development of proteolytic activity and protein degradation during the germination of *Pisum sativum* L. In: *Planta*, 1975, vol.124, p. 77-87. ISSN 0032-0935. doi:10.1007/BF00390070.
- 95. BAUMGARTNER, B., CHRISPEELS, M.J. Purification and characterization of vicilin peptidohydrolase, the major endopeptidase in the cotyledons of mung-bean seedlings. In: *Eur. J. Biochem.*, 1977, vol.77, p. 223-233. ISSN 0014-2956. <u>doi:org/10.1111/j.1432-1033.1977.tb11661.x</u>
- RUDAKOVA, A.S., CHERDIVARA, A.M., WILSON, K.A., SHUTOV A.D. Seed storage globulins: origin and evolution of primary and higher-order structures. In: *Biochemistry* (Moscow), 2015, vol.80, p.1354-1361. ISSN 0006-2979. doi:10.1134/S000629791510017X
- 97. LEHMANN, R., MACHNÉ, R., HERZEL, H. The structural code of cyanobacteria genomes. In: Nucleic Acids Res., 2014, vol.42, p. 8873-8883. ISSN 0305-1048. doi.org/10.1093/nar/gku641
- 98. SHUTOV, A.D., BRAUN, H., CHESNOKOV, Yu.V., BÄUMLEIN, H. A gene encoding a vicilin-like protein is specifically expressed in fern spores. Evolutionary pathway of seed storage globulins. In: *Eur. J. Biochem.*, 1998, vol.252, p. 79-89.ISSN 0014-2956.
- 99. SCHALLAU, A. et al. Phylogenetic footprints in fern spore- and seed-specific gene promoters. In: *Plant J.*, 2008, vol.53, p. 414-424. ISSN 0960-7412 doi:10.1111/j.1365-313X.2007.03354.x
- 100.KAKHOVSKAYA, I.A., RUDAKOVA, A.S., MANTEUFFEL, R. Legumin-like and vicilinlike proteins from spores of the fern *Matteuccia struthiopteris*. In: *Journal of Plant Physiology*, 2003, vol.160, p. 583-588. ISSN 0176-1617. doi:10.1078/0176-1617-00976
- 101. **КЕРДИВАРЭ, А.**, КАХОВСКАЯ, И., ШУТОВ, А. Эволюционные предшественники запасных глобулинов семян в геномах зеленых водорослей и бактерий. Conferința științifică națională consacrata jubileului de 90 de ani din ziua nașterii academicianului B. Melnic. USM, 12 februarie 2018, p. 222-225. ISBN 978-9975-71-971-1.
- 102.SHUTOV, A.D., BLATTNER, F.R., BÄUMLEIN, H., MÜNTZ, K. Storage and mobilization as antagonistic functional constraints of seed storage globulin evolution. In: J. Exp. Bot., 2003, vol.54, p. 1645-1654. ISSN 0022-0957. doi:10.1093/jxb/erg165

- 103.TAN-WILSON, A.L., WILSON, K.A. Mobilization of seed protein reserves. In *Physiol. Plant.*, 2012, vol.145, p. 140-153. ISSN 0031-9317. doi:10.1111/j.1399-3054.2011.01535.x
- 104.VAINTRAUB, I.A., MORARI, D. Applying the increase in rate constants of cooperative proteolysis to the determination of transition curves of protein denaturation. In: *Biochem. Biophys. Methods*, 2003, vol.57, p. 191-201. ISSN 0165-022X. doi:10.1016/s0165-022x(03)00106-<u>4</u>
- 105.SCHECHTER, I., BERGER, A. On the size of the active site in proteases. I. Papain. In: Biochem. Biophys. Res. Commun., 1967, vol.27, p. 157-162. ISSN 0006-291X. doi:10.1016/s0006-291x(67)80055-x
- 106.LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. In: *Nature*, 1970, vol.227, p. 680-685. ISSN 0028-0836.
- 107.VAINTRAUB, I.A., YATTARA, H.B. Proteolysis of Kunitz soybean inhibitor. Influence on its activity. In: *J. Agric. Food Chem.*, 1995, vol.43, p. 862-868. ISSN 0021-8561.
- 108.JIN, T., CHEN, Y-W., HOWARD, A., ZHANG, Y-Z. Purification, crystallization and initial crystallographic characterization of the *Ginkgo biloba* 11S seed globulin ginnacin. In: <u>Acta</u> <u>Crystallogr Sect F</u>, 2008, vol.64, p. 641-644. ISSN 2053-230X. doi:10.1107/S1744309108016242
- 109.HÄGER, K.P. et al. Evolution of seed storage protein genes: legumin genes of *Ginkgo biloba*. In: J. Mol. Evol., 1995, vol.41, p. 457-466. ISSN 0022-2844. doi:10.1007/bf00160317
- 110.CHERDIVARA, A.M. et al. Limited proteolysis regulates massive degradation of ginnacin, storage 11S globulin from ginkgo seeds. In: *Molecular Life* [online], Romania, 2018, vol. 2, p. 2-6.ISSN 2559-611X. doi:10.26600/MolLife.1.2.1.2018
- 111.КЕРДИВАРЭ, А. и др. Роль ограниченного протеолиза в инициации массированной деградации запасных 11S глобулинов семян. În: Materialele Conferinței ştiințifice Internaționale "Genetica, fiziologia şi ameliorarea palntelor". (Ediția a VI-a), AŞM, Chişinau, 9-10 octombrie 2017, p.26-29. ISBN 978-9975-56-46-36-2
- 112.WILSON, K.A, TAN-WILSON, A. Proteolysis of the peanut allergen Ara h1 by an endogenous aspartic proteinase. In: *Plant Physiol. Biochem.*, 2015, vol.96, p. 301-310. ISSN 0981-9428. doi:10.1016/j.plaphy.2015.08.008
- 113. CHERDIVARĂ, A., RUDAKOVA, A., ȘUTOV A. Globulinele de rezervă 7S din semințe ca alergeni. Imunoreactivitatea încrucișată între globulinele de rezervă 7S și 11. În: *Studia Universitatis*, seria *Științe reale și ale naturii*. 2016, nr.1 (91), p.56-60. ISSN 1814-3237
- 114. CHERDIVARĂ, A., RUDAKOVA, A., RUDAKOV, S., ŞUTOV, A. Proteoliza limitată a alergenului Ara h1, globulina de rezervă 7S din semințele de arahide. În: *Studia Universitatis*, seria *Ştiințe reale și ale naturii*. 2016, nr.6(96), p.10-15. ISSN 1814-3237.
- 115.CHERDIVARA, A. Limited proteolysis as a means to reduce the allergenicity of seed storage globulins. In: *Agricultural Biology*, 2018, vol. 53, p. 475-484. ISSN 2412-0324. doi:10.15389/agrobiology.2018.3.475eng
- 116. CHERDIVARĂ, A., RUDAKOVA, A., RUDAKOV, S., ŞUTOV, A. Alergenul Ara h3, globulina de rezervă din semințele de arahide. 1. Proteoliza limitată cu papaină. În: *Studia Universitatis*, seria *Ştiințe reale și ale naturii*. 2017, nr.1(101), p.37-40. ISSN 1814-3237
- 117. CHERDIVARĂ, A., RUDAKOVA, A., ŞUTOV, A. Globulinele de rezervă 11S ca alergeni.
  În: Studia Universitatis, seria Științe reale și ale naturii. 2015, nr.6 (86), p.9-11. ISSN 1814-3237

- 118. **CHERDIVARĂ, A.**, RUDAKOVA, A., RUDAKOV, S., ŞUTOV, A. Alergenul Ara h3, globulina de rezervă din semințele de arahide. 2. Proteoliza limitată cu tripsină. În: *Studia Universitatis*, seria *Științe reale și ale naturii*. 2017, nr.1(101), p.41-45. ISSN 1814-3237
- 119.IVANCIUK, O. et al. The property distance index PD predicts peptides that cross-react with IgE antibodies. In: *Mol. Immunol.*, 2009, vol.46, p. 873-883. ISSN 0161-5890.
- 120.CHERDIVARĂ, A. Imunoreactivitatea încrucişată potențială a globulinelor de rezervă din semințe. În: *Studia Universitatis*, seria *Științe reale și ale naturii*. 2017, nr.6 (106), p. 19-23. ISSN 1814-3237

### ANEXE

### Anexa 1. Act de inplementare a rezultatelor științifice în instruire

MINISTERUL EDUCAȚIEI, CULTURII ȘI CERCETĂRII AL REPUBLICII MOLDOVA

> UNIVERSITATEA DE STAT DIN MOLDOVA

MD-2009, mun. Chişinău str. A.Mateevici 60 tel.: (+373) 22244821, fax: 22244248 www.usm.md, email: rector@usm.md



MINISTRY OF EDUCATION, CULTURE AND RESEARCH OF THE REPUBLIC OF MOLDOVA

### MOLDOVA STATE UNIVERSITY

MD-2009, Chisinau A.Mateevici str. 60 phone: (+373) 22244821, fax: 22244248 www.usm.md, email: rector@usm.md

21.05.20 Nr. 01/560

### ACT DE IMPLEMENTARE

### a rezultatelor tezei de doctor

### "Globulinele de rezervă 7S și 11S din semințe: evoluție, proteoliză, alergenicitate"

Prin prezenta se confirmă că rezultatele științifice, efectuate în cadrul tezei de doctor în biologie cu tema "Globulinele de rezervă 7S și 11S din semințe: evoluție, proteoliză, alergenicitate" realizată de dna Cherdivară Ala sunt implementate în curriculumul universitar pentru studii superioare de licență (ciclul I) la disciplina *Biochimie* și de master (ciclul II) la disciplina *Proteomica* la Departamentul Biologie și Ecologie a Facultății Biologie și Pedologie, Universitatea de Stat din Moldova.

Decanul Facultății de Biologie și Pedologi Dr., conf. M. LESANU mnătura Prorector pentru activitatea didactica US Dr., hab.,prof. univ. O. DANDARA semnătura

# DECLARAȚIA PRIVIND ASUMAREA RĂSPUNDERII

Subsemnata, Cherdivară Ala, declar pe răspundere personală că materialele prezentate în teza de doctorat sunt rezultatul propriilor cercetări și realizări științifice. Conștientizez că, în caz contrar, urmează să suport consecințele în conformitate cu legislația în vigoare.

03.09.19

Cherdivara Ala

# CV –ul AUTORULUI

# CHERDIVARĂ Ala

# **Data nașterii**: 10.11.1975

**Locul nașterii**: s. Bilicenii Vechi, r. Sângerei, R. Moldova

### Cetățenie: MD

### Studii:

990-1992	-studii liceale, liceul teoretic nr.2, Chișinău
992-1997	-studii de licență, Universitatea de Stat din Moldova, Facultatea Biologie și Pedologie, specialitatea Biologie
997-2000	-studii de doctorat, Universitatea de Stat din Moldova, specialitatea 163.02- Biochimie

### Activitatea profesională:

septembrie 2000 - prezent -	lector universitar, departamentul Biologie și Ecologie,
	facultatea Biologie și Pedologie, Universitatea de Stat din
	Moldova
ianuarie 2015 – prezent –	cercetător științific, LCȘ Biochimia Plantelor, facultatea
	Biologie și Pedologie, Universitatea de Stat din Moldova

Domeniile de interes științific: biochimie, biologie moleculară

### Participări la foruri științifice (nationale și internaționale)

### 2015

- Conferința științifică națională cu participare internațională *Integrare prin cercetare și inovare*. USM, Moldova, Chișinau, 10-11 noiembrie

### 2016

- Conferința științifică națională cu participare internațională *Integrare prin cercetare și inovare*. USM, Moldova, Chișinau, 28-29 septembrie
- Всероссийская конференция с международным участием Агроэкосистемы в естественных и регулируемых условиях: от теоретической модели к практике прецизионного управления. Россия, Санкт-Петербург, 21-23 сентября

### 2017

- Международная научная конференция *Тенденции развития агрофизики: от современных проблем земледелия и растениеводства к технологиям будущего.* Россия, Санкт-Петербург, 27-29 сентября
- Conferința științifică internaționale *Genetica, fiziologia și ameliorarea plantelor*. AȘM, Moldova, Chișinau, 9-10 octombrie

# 2018

- Conferința științifică națională consacrata jubileului de 90 de ani din ziua nașterii academicianului B. Melnic. USM, Moldova, Chișinau; 12 februarie
- The International Congress on Oil and Protein Crops. Meeting of the EUCARPIA Oil and Protein Crops Section, Moldova, Chisinau, May 20-24

# Participări la proiecte științifice naționale și internaționale:

### - proiect instituțional

15.817.05.02F "Substanțele biologic active ca bază a valorificării biotehnologiilor moderne în modularea și adaptarea proceselor metabolice ale organismelor vii", 2015-2019.

### Lucrări științifice și științifico-metodice publicate:

18 lucrări științifice: 3 articole în reviste internaționale cu factor de impact până la 1,75,

6 articole în revista Studia Universitatis, USM

9 comunicări la foruri naționale și internaționale.

Apartenența la societăți: Societatea de Biochimie din Moldova.

Cunoașterea limbilor: limba română – limba maternă, limba rusă – bine, limba franceză – mediu

### Date de contact:

LCŞ Biochimia Plantelor Universitatea de Stat din Moldova str. A.Mateevici 60, MD-2009, Chișinau, e-mail: <u>alacherdivara@mail.ru</u> tel: +37369920027