

Limited proteolysis initiates massive degradation of ginnacin, storage 11S globulin from *Ginkgo* seeds

Alla Cherdivară¹, Angela Rudakova¹, Tatiana Stepurina¹, Sergei Rudakov¹, Karl Wilson², Andrei Shutov^{1*}

¹State University of Moldova, MD-2009 Chişinău, Moldova

²State University of New York at Binghamton, NY 13902-6000 Binghamton, NY, USA

(Received 24 September, 2017; accepted 12 January 2018)

Abstract

Limited proteolysis of α -chains, followed by a deep cleavage of $\alpha\beta$ -subunits *via* one-by-one mechanism, is observed in the course of papain hydrolysis of ginnacin, the storage 11S globulin from *Ginkgo biloba* seeds. An independent analysis of kinetics of the limited and one-by-one proteolyses allows suggesting that the availability of ginnacin for complete degradation is acquired due to loosening of the inter-subunit interactions in its oligomeric molecule. Thus, the limited proteolysis plays a key role in the subsequent massive degradation of ginnacin. A similar succession of the reactions of limited and one-by-one proteolyses was observed previously during investigation of soybean and pumpkin 11S globulin proteolysis. These data, as well as the results of computer analysis of 11S globulin crystal structures, allow suggesting that the key role of limited proteolysis, which appeared in the early stages of 11S globulin evolution, has been inherited by other storage proteins from 11S globulin family.

Keywords: *Ginkgo biloba*, seed storage globulin, limited proteolysis, papain, kinetics

INTRODUCTION

11S globulins and 7S globulins, the most characteristic seed storage proteins, are degraded during seed germination to nourish growing seedlings. During the evolution of seed storage globulins, extended disordered inserts of enhanced susceptibility to proteolytic attack (1) appeared in their sequences that have been inherited from spore plants (2). These inserts determine the susceptibility of the globulins to limited proteolysis that occurs at the very beginning of their degradation during seed germination and *in vitro* (1, 2). Further deep splitting of 11S and 7S globulins follows the one-by-one mechanism (3) that calls forth a massive degradation of their molecules (1). The native soybean (4) and pumpkin (5) 11S globulins, and common bean 7S globulin (6) are initially inaccessible to the one-by-one proteolysis during proteinase action. Only after completion of limited proteolysis, resulting in regular alterations of their

native structure, do soybean, pumpkin and common bean globulins acquire the ability for degradation by the one-by-one mechanism (4-6). Thus, the initial limited proteolysis plays a regulatory role, delaying the onset of the massive degradation of these storage proteins.

Does the regulatory role of the limited proteolysis of 11S globulins appear in such an early evolutionary level as *Ginkgophyta*? It seems probable because the primary structures of seed 11S globulins, including ginnacin, a storage 11S globulin from *Ginkgo* seeds (7), are highly conserved (1, 8). Similar to other storage 11S globulins (1), the ginnacin subunits are synthesized as a single polypeptide chain proteolytically cleaved into N- and C-terminal domains (α - and β -chains, respectively) conjoined with a disulfide bond (7).

To check the above hypothesis, the kinetics of proteolysis of ginnacin were studied in this work. In addition, all crystal structures of 11S globulins hitherto

*Corresponding author: Moldova, MD-2009 Chişinău, Mateevici Str. 60, State University of Moldova, Fax: (37)322-244+248; Email: shutovandrei@yahoo.com

available were analyzed in comparison with a model structure of ginnacin.

MATERIALS AND METHODS

Isolation and purification of ginnacin was carried out essentially as described earlier (7, 9). The meal from ginkgo (*Ginkgo biloba* L.) seeds was extracted (1:10, w/v) with 1 M NaCl in Buffer A (0.02 M Tris-HCl, pH 8.0, 0.02% NaN_3 , 1 mM EDTA). Ginnacin was salted out from the extract with the addition of ammonium sulfate up to 60% saturation. The precipitate was dissolved in 3.0 M NaCl in Buffer A, and the solution was loaded onto a column of Phenyl Sepharose CL-4B (Pharmacia Biotech, Sweden) equilibrated with 3.0 M NaCl in Buffer A. After removal of the non-adsorbed fraction, ginnacin was eluted with 2.0 M NaCl in Buffer A. The obtained preparation was additionally purified by gel filtration using a column (0.9 x 70 cm) of Sephacryl S-300 (Pharmacia Biotech, Sweden) equilibrated with 1 M NaCl in Buffer A.

Proteolysis of ginnacin. For the ginnacin digestion, the reaction mixture containing 5 mg/mL substrate and 100 $\mu\text{g/mL}$ papain (Sigma Life Science, USA) in Buffer B (0.037 M phosphate/citrate buffer, pH 5.6, 1 M NaCl, 0.02% NaN_3 , 1 mM EDTA, 10 mM 2-mercaptoethanol) was incubated for 24 h at 30° C. To follow the time course of the digestion, aliquots of the incubation mixture were sampled from time to time and the reaction was stopped by addition of trichloroacetic acid (TCA) to a final concentration of 5% (w/v). In some cases, the reaction was stopped by addition of E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino) butane (Sigma, Life Science, USA) to the final concentration of 10 μM . The obtained protein samples were subjected to gel filtration using Sephacryl S-300 column and analyzed by electrophoresis on polyacrylamide gel in the presence of sodium dodecyl sulphate (SDS-PAGE).

Electrophoresis and analysis of electropherograms. SDS-PAGE was carried out in 15% (w/v) gels using the buffer system of Laemmli (10). Prior to electrophoresis, the residual TCA in precipitated protein samples was removed by washing with acetone, the pellets were dried and dissolved in sample buffer with 2-mercaptoethanol under standard conditions. Molecular weight markers from Bio-Rad (USA) were used. The electropherograms were scanned (ImageScanner III, GE Healthcare) and analyzed quantitatively using Phoretix 1D Gel Analysis v.5.10 software. The relative molar amounts of polypeptides detected by SDS-PAGE in the intact ginnacin and in the residual protein after proteolysis were calculated using their molecular masses and relative weight concentrations.

Analysis of proteolysis kinetics. For analysis of proteolysis kinetics, residual TCA-insoluble protein was determined by a dye-binding method (11). The

number-average molecular masses of the intact ginnacin subunits and of those of the residual protein samples in the course of the reaction were calculated using the molecular masses of polypeptides separated by SDS-PAGE and their relative weight amounts. Concentrations of the residual protein P_t and its molecular masses M_t were expressed as percentages of the initial values P^0 and M^0 .

The analysis of proteolysis kinetics was carried out following our previously described strategy (4, 5). During mixed-type proteolysis, when limited and one-by-one proteolyses occur either in parallel independently from each other or successively, the input of the limited proteolysis in the total decline of the protein weight concentration of the hydrolyzate ($P^0 \rightarrow P^t$) is equal to $M^0 - M^t$. Hence, the equation $P^t + (M^0 - M^t) = f(t)$ exclusively describes the kinetics of the one-by-one proteolysis, which occurs as a pseudo-first order reaction (12, 13). Accordingly, the decline of the relative molecular mass value of the protein substrate ($M^0 \rightarrow M^t$) exclusively describes the kinetics of the limited proteolysis.

Modeling of protein structures and computer analyses. Homology modeling of tertiary and quaternary structure of ginnacin was carried out with the Swiss-Model program (14, 15) using the crystal structure pdb|3qac (11S globulin from *Amaranthus hypochondriacus*) as a template. The same program was used for modeling of quaternary structures of those 11S globulins in which only tertiary structures of A subunit were available (pdb|2e9q, pdb|3c3v, pdb|3qac). The SSPRO 8 (Scratch Protein Predictor) and DisEMBL 1.5 programs were also used for prediction of secondary structures and disordered regions, respectively, in ginnacin tertiary structure.

The surface area of an amino acid residue accessible to the solvent (ASA) in the crystal and model structures of 11S globulin oligomers, expressed as Å^2 , was calculated using the program <http://cib.cf.ocha.ac.jp/bitool/ASA/>. The relative ASA value of an X residue in 11S globulins was expressed in percent of its accessibility in the GXG tripeptide. The DeepView/Swiss-PdbViewer v.3.7 program was used for contouring of ribbon diagrams, identification of the inter-subunit hydrogen bonds in 11S globulin oligomers and calculation of a root-mean-square deviation (RMSD) of superimposed subunit structures of ginnacin (model) and other 11S globulins (crystal structures).

RESULTS AND DISCUSSION

According to SDS-PAGE data, the subunits of native ginnacin are composed of at least two variants of α -chains (28.3 kD and 26.7 kD) and β -chains (21.1 kD). Thus, the hexamer structure of the ginnacin molecule (9) contains at least two kinds of $\alpha\beta$ subunits (Fig. 1, lane 1). Similarly to other 11S globulins (16), the ginnacin α - and β -chains

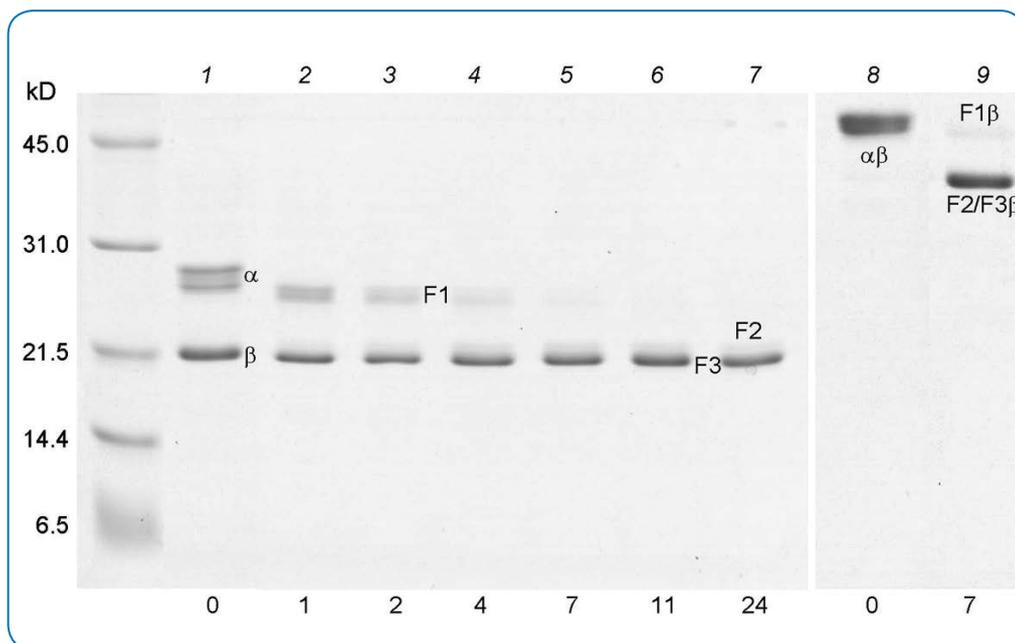


Fig. 1. SDS-PAGE of intact ginnacin and products of its proteolysis in the presence (lanes 1-7) and in the absence (lanes 8, 9) of 2-mercaptoethanol. Masses of molecular markers are shown on the left side. Numbers at the bottom correspond to the reaction time (h).

are conjoined with a disulfide bond (lanes 1 and 8). A single amino acid sequence (caa90641) of ginnacin subunits comprising α - (27.5 kD) and β - (21.1 kD) chains is known.

Proteolysis of ginnacin starts with the formation of at least two intermediate α -chain fragments F1 with molecular masses of 25.8 kD and 25.2 kD (Fig. 1, lanes 2-5). During further proteolysis, the fragments F1 disappear, and two fragments appear: the major fragment F3 (21.1 kD), which has a mobility equal to that of β -chains, and the minor fragment F2 (22.0 kD). Relatively low-molecular-mass (less than 20 kD) TCA-insoluble products are absent. Thus, the β -chains of ginnacin are inaccessible to limited proteolysis, which is characteristic of the majority of 11S globulins studied (1), and the

hydrolysis of α -chains consists of their successive shortening.

During gel filtration on Sephacryl S-300 column of the ginnacin proteolysis products formed after 24 h of the reaction, a single peak of a high-molecular-mass protein was observed. Its SDS-PAGE pattern was found the same as shown in Fig. 1, lane 7.

In the first proteolysis stage (up to 2 h), the relative decline of the protein weight concentration P does not exceed the decline of its molecular mass M (Fig. 2, plots 1 and 2). This evidences the absence of any detectable indications on the one-by-one proteolysis in the first stage. These indications appear at the intermediate stage (2-7 h), and further proteolysis occurs as a pseudo-first order reaction (Fig. 2, plot 3). The extrapolation of the linear part of

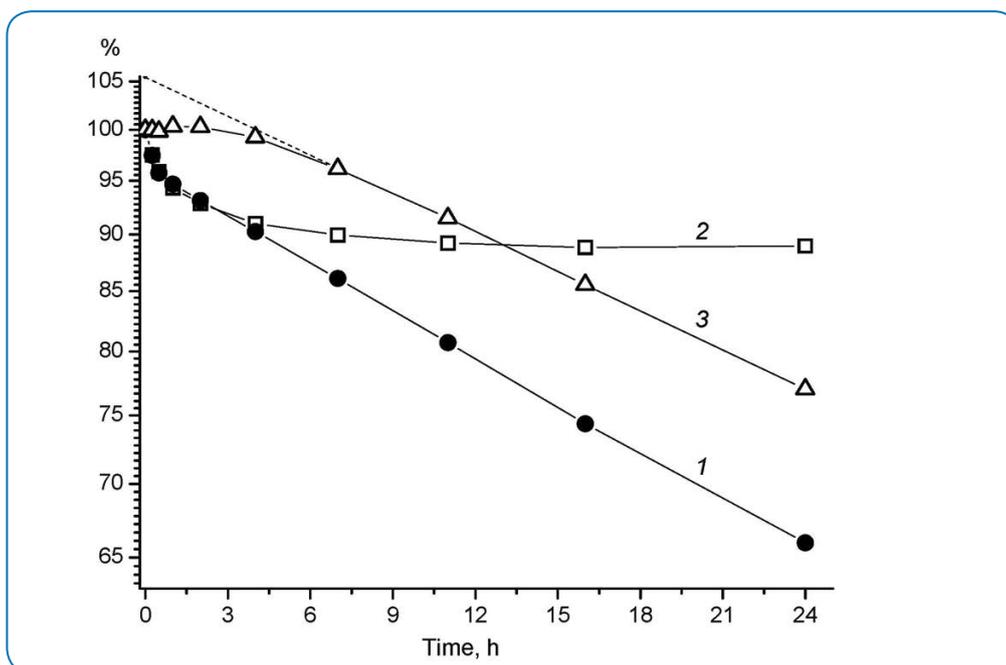


Fig. 2. The proteolysis kinetics of ginnacin. Plot 1, weight concentration P of the residual (TCA-insoluble) protein, and plot 2, its number-average molecular mass M . P and M are expressed as percentage of the initial values. Plot 3 ($P\% + [100\% - M\%]$) describes the kinetics of exclusive one-by-one proteolysis (see above). The ordinate axis is logarithmic.

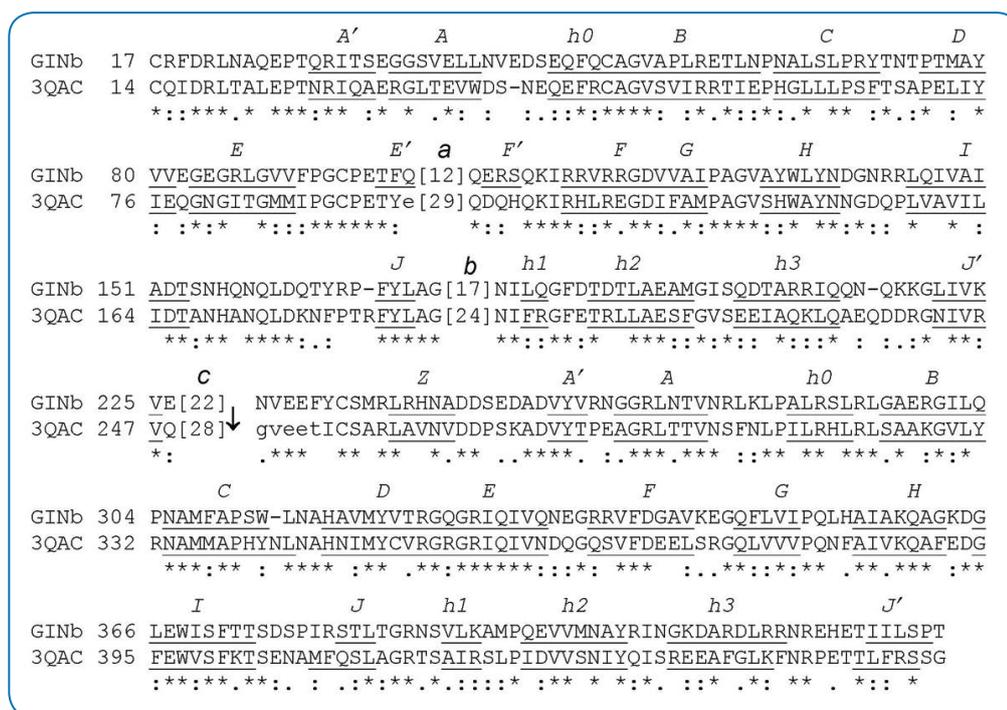


Fig. 3. Aligned subunit sequences of ginnacin (caa90641, GINb) and 11S globulin from *Amaranthus hypochondriacus* (pdb|3qac). Secondary structures of both pdb|3qac (template) and ginnacin (model) are shown. The number of missing residues within variable regions *a*, *b* and *c* are shown in square brackets. An arrow indicates the border between α - and β -chains (N- and C-terminal domains, respectively).

plot 3 to zero time led to a P value that exceeds 100%. This fact points to the dependence of the rate constant of one-by-one proteolysis of ginnacin on the structural alterations resulting from the limited proteolysis.

Only initial information on the crystal structure of ginnacin is currently available (9). Therefore, structural alterations of ginnacin due to limited proteolysis can be estimated only in the first approximation on the basis of its model tertiary and quaternary structures as well as on the basis of available information concerning proteolysis regularities of other seed 11S globulins. All the secondary structures characteristic of 11S globulins (16) are present

in the structural model of ginnacin (Fig. 3). Three extended disordered regions in the 11S globulin α -chains containing amino acid residues with enhanced accessibility to the solvent are potentially susceptible to papain attack (1): *a*) between β -strands *E'* and *F'*; *b*) between β -barrel and the group of α -helices; *c*) C-terminal region.

The results of homology modeling of ginnacin tertiary structure are supported by alternative methods (SSPRO 8 and DisEMBL 1.5) tested. Additionally, the results of RMSD measurements also support the validity of the ginnacin structural model. Thus, the RMSD of 0.32 Å was obtained for 368 C α atoms (85%) of superimposed

Table 1. The accessibility to the solvent and the inter-subunit hydrogen bonds in 11S globulin oligomers

Species	pdb	Structural region			
		ASA, Å ²	H-bonds		
		<i>B - E</i>	<i>h1 - J'</i>	<i>A' - J</i>	<i>J - J'</i>
<i>Ginkgo biloba</i>		15.3	42.4	10	11
<i>Amaranthus hypochondriacus</i>	3qac	11.6	39.7	14	8
<i>Prunus dulcis</i>	3fz3	13.7	33.2	16	10
<i>Brassica napus</i>	3kgl	13.5	36.8	15	12
<i>Cucurbita maxima</i>	2e9q	12.9	33.9	11	11
<i>Glycine max</i>	2d5h	13.4	34.7	9	12
<i>Glycine max</i>	1fxz	15.2	34.8	11	9
<i>Pisum sativum</i>	3ksc	13.0	37.6	14	9
<i>Arachis hypogaea</i>	3c3v	15.9	37.0	16	10
<i>Average values</i>		13.8	36.7	13	10

ASA=average accessible surface area calculated per one amino acid residue in the N-terminal half of the β -barrel (β -strands *B - E*) and in the region of β -helices *h1*, *h2*, *h3* and β -strand *J'*. H-bonds=the number of hydrogen bonds formed by amino acid residues in the region between β -strands *A'* and *J*, and in the C-terminal region between β -strands *J* and *J'*.

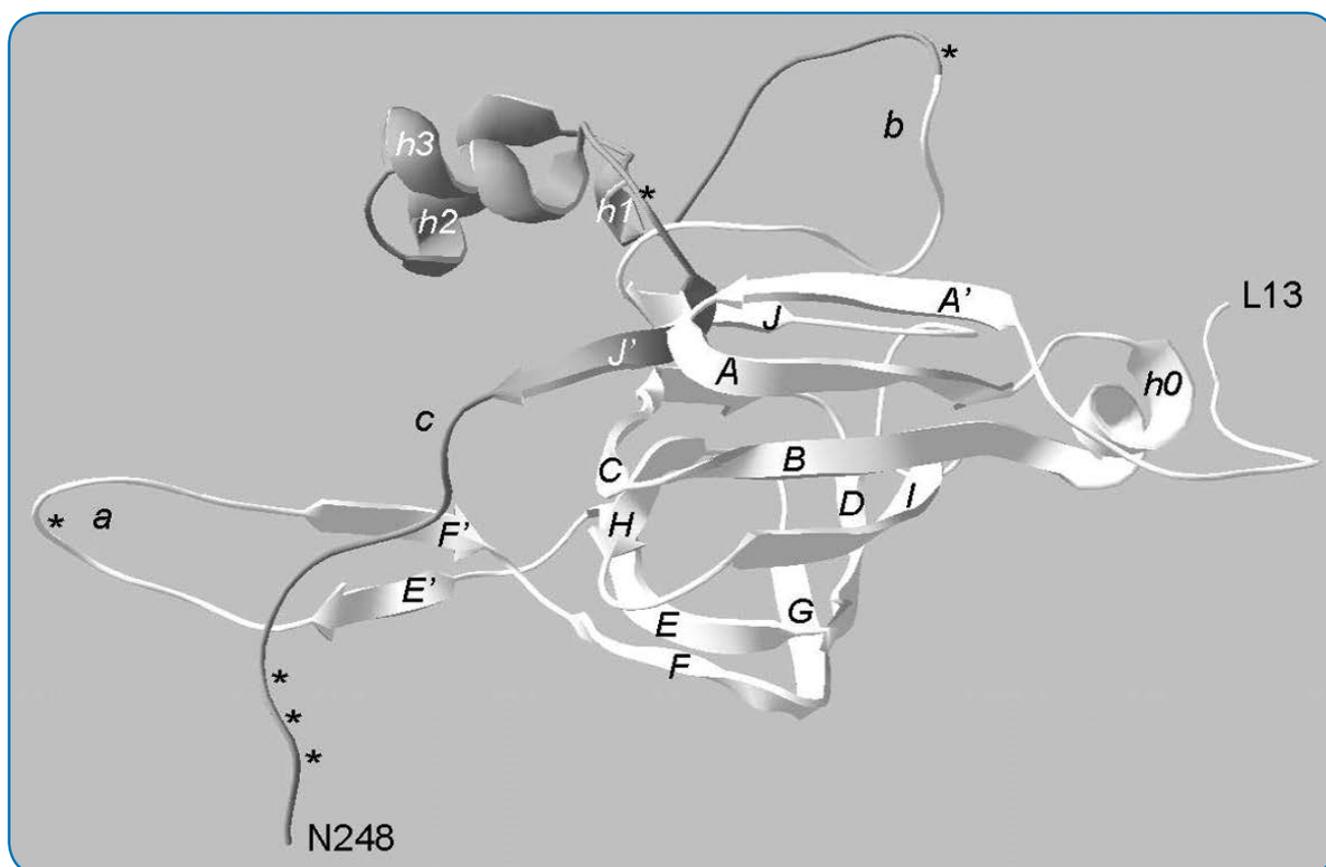


Fig. 4. Ribbon diagram of a model tertiary structure of ginnacin α -chain. Similarly to other 11S globulins (16), a β -barrel of anti-parallel β -strands BCDEFGHI, conjoined with a group of α -helices $h1$, $h2$, $h3$, and extended by additional β -strands A-A', E'-F' and J-J', and α -helix $h0$, represents a structural basis of the ginnacin α -chain. The regions a , b and c are potentially susceptible to limited proteolysis. Positions of amino acid residues of enhanced (>90%) accessibility to the solvent, that correspond to papain substrate specificity (17), are shown by asterisks. The dark part of the diagram corresponds to the C-terminal sequence region that is presumably detached after completion of limited proteolysis.

ginnacin and 3qac structures. In average, the RMSD of 0.98 ± 0.07 Å was obtained for 352 ± 6 C α atoms (81%) of superimposed ginnacin structure and all crystal structures of other 11S globulins hitherto available (Table 1).

The loop a between β -strands E' and F' in ginnacin structure is inaccessible to papain action because no fragments of molecular mass lower than 20 kD corresponding to the N-terminal half of the β -barrel (β -strands A'-E', Fig. 4) were detected among proteolysis products (Fig. 1). Similarly, this loop is inaccessible to papain attack in region soybean (4) and pumpkin (5) 11S globulins.

It seems likely that the initial limited proteolysis of ginnacin, which consists of the detachment of the C-terminal α -chain region most susceptible to proteolysis in 11S globulins (1), is further finalized by the detachment of the region of α -chain helices $h1$ - $h3$ together with the β -strand J' (Fig. 4); this region is destroyed up to the formation of low-molecular mass products non-detectable by SDS-PAGE. The destruction of this region during papain hydrolysis of soybean (4) and pumpkin (5) 11S globulins (as well as of ginnacin) produces regular alterations of their native structure.

As a result, these proteins acquire susceptibility to the one-by-one proteolysis.

Amino acid sequences of 11S globulins outside the limits of variable regions a , b and c (Figures 3, 4) are highly conserved, including the structural region of α -helices $h1$ - $h3$ and β -strand J' (16). Possibly, this region is capable of destruction during proteolysis not only of *Ginkgo*, soybean and pumpkin 11S globulins but also of other proteins from this family.

To check this hypothesis, the amino acid sequences of α -chains from all 11S globulins of known tertiary structures and also from ginnacin were compared. As follows from Table 1, solvent accessibility of the region of β -helices $h1$ - $h3$ and β -strand J' is 2.3-3.4 times higher than that of the β -barrel N-terminal half (β -strands B-E), which is not susceptible to limited proteolysis (2). This highlights the potential ability of this region for a deep splitting during proteolysis of 11S globulins.

Identification of inter-subunit hydrogen bonds in the quaternary structures of 11S globulins performed in this work and previously (4, 18) revealed that up to a half of these bonds were formed with the participation of amino acid residues inside the C-terminal region of the

subunit A out from limits of the β -strand *J*. It seems likely that these bonds in soybean 11S globulin (4) as well as in other 11S globulins are disrupted due to destruction of the region α -helix *h1* - β -strand *J'*. As a result, the inter-subunit interactions become essentially loosened, and the 11S globulins acquire the ability for a reversible dissociation that exposes peptide bonds, which are potentially susceptible for proteolysis but are masked in the native molecules (2).

The results of the present work indicate that the susceptibility of 11S globulins to a massive degradation by

the one-by-one mechanism, which is acquired only after completion of limited proteolysis, appeared in the early stages of evolution of these highly conserved proteins (1, 16). Therefore, it is possible to suggest that the regulatory role of the limited proteolysis was preserved in the course of evolution of many 11S globulins.

Author contributions: experimental work: A. Cherdivarä, T. Stepurina, A. Rudakova; analysis: S. Rudakov; writing: K. Wilson, A. Shutov.

Conflicts of interest: None declared.

REFERENCES

1. Shutov AD, Wilson KA, *Seed storage globulins: their descent from bacterial ancestors and mechanisms of degradation*, In: **Globulins: Biochemistry, Production and Role in Immunity**, Milford SD, ed, Nova Science Publishers, New York, 2014, pp 71-104.
2. Rudakova AS, Cherdivarä AM, Wilson KA, Shutov AD, *Seed storage globulins: origin and evolution of primary and higher order structures*, **Biochemistry (Moscow)**, **80**, 1354-1361 (2015).
3. Rupley JA, *Susceptibility to attack by proteolytic enzymes*, **Methods Enzymol.**, **11**, 905-917 (1967).
4. Shutov A, Rudakova A, Rudakov S, Kakhovskaya I, Schallau A, Maruyama N, et al, *Limited proteolysis regulates massive degradation of glycinin, storage 11S globulin from soybean seeds: An in vitro model*, **J. Plant Physiol.**, **169**, 1227-1233 (2012).
5. Rudakova AS, Rudakov SV, Kakhovskaya IA, Shutov AD, *11S storage globulin from pumpkin seeds: regularities of proteolysis by papain*, **Biochemistry (Moscow)**, **79**, 820-825 (2014).
6. Senyuk V, Rotari V, Becker C, Zakharov A, Horstmann C, Müntz K, Vaintraub I, *Does an asparaginyl-specific cysteine endopeptidase trigger phaseolin degradation in cotyledons of kidney bean?* **Eur. J. Biochem.**, **258**, 546-558 (1998).
7. Häger KP, Braun H, Czihal A, Müller B, Bäumlein H, *Evolution of seed storage protein genes: legumin genes of Ginkgo biloba*, **J. Mol. Evol.**, **41**, 457-466 (1995).
8. Shutov AD, Kakhovskaya IA, Braun H, Bäumlein H, Müntz K, *Legumin and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene*, **J. Mol. Evol.**, **41**, 1057-1069 (1995).
9. Jin T, Chen YW, Howard A, Zhang YZ, *Purification, crystallization and initial crystallographic characterization of the Ginkgo biloba 11S seed globulin ginnacin*, **Acta Cryst.**, **F64**, 641-644 (2008).
10. Laemmli UK, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*, **Nature**, **227**, 680-685 (1970).
11. Vaintraub IA, Yattara HB, *Proteolysis of Kunitz soybean inhibitor. Influence on its activity*, **J. Agric. Food Chem.**, **43**, 862-868 (1995).
12. Shutov AD, Pineda J, Senyuk VI, Reva VA, Vaintraub IA, *Action of trypsin on soybean glycinin. Mixed-type proteolysis and its kinetics; molecular mass of glycinin-T*, **Eur. J. Biochem.**, **199**, 539-543 (1991).
13. Vaintraub IA, *Kinetics of co-operative proteolysis*, **Nahrung**, **42**, 59-60 (1998).
14. Arnold K, Bordoli L, Kopp J, Schwede T, *The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling*, **Bioinformatics**, **22**, 195-201 (2006).
15. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T, *SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information*, **Nucleic Acids Research**, **42**, W252- W258 (2014).
16. Tandang-Silvas MRG, Fukuda T, Fukuda C, Prak K, Cabanos C, Kimura A, Itoh T, Mikami B, Utsumi S, Maruyama N, *Conservation and divergence on plant seed 11S globulins based on crystal structures*, **Biochim. Biophys. Acta**, **1804**, 1432-1442 (2010).
17. Schechter I, Berger A, *On the size of the active site in proteases. I. Papain*, **Biochem. Biophys. Res. Commun.**, **27**, 157-162 (1967).
18. Adachi M, Takenaka Y, Gidamis AB, Mikami B, Utsumi S, *Crystal structure of soybean proglycinin A1aB1b homotrimer*, **J. Mol. Biol.**, **305**, 291-305 (2001).