Stabilization of enzyme proteins using knowledge on thermophiles and thermophile enzymes

TAIRO OSHIMA

Introduction

Thermophiles are interesting microorganisms. Their biopolymers and cell constituents, that is, proteins, enzymes, nucleic acids, ribosomes, membranes, and viruses, are resistant to heat without any exception. They live in thermal environment, both natural and artificial such as hot springs, thermal vents on the deep-sea bottom, or industrial hot water reservoirs and wastes [1,2].

An extreme thermophile, *Thermus thermophilus,* was isolated from a hot spring in Izu peninsula, Japan, many years ago [3]. This bacterium is able to grow up to 85 °C, and used widely in biochemical studies. A taxonomically closely related species, *T. aquaticus* is also used by many investigators. Recently hyperthermophiles were isolated and are used in some laboratories. Hyperthermophiles are microorganisms capable of growing at 90 °C or higher temperatures. At moment, the highest record for growth is 115 °C [4].

T. thermophilus is a gram negative, aerobic, non-spore forming, rod cell, and nothing is unique in the cell structure compared with that of the mesophiles such as *Escherichia coli.* However, the cell structure is more stable, for instance, the spheroplasts can be washed with distilled water. Likewise enzymes of this extreme thermophile have been shown to be highly homologous

Table 1. Thermal denaturation temperatures of biopolymers.
Tabuľka 1. Teploty termálnej denaturácie biopolymérov.

BIOPOLYMER [1]	SOURCE	
	Mesophilic [2]	*T. thermophilus* [3]
IPMDH	50	90
PGK	58	100
tRNA	78	88

1 - biopolymér, 2 - mezofilný mikroorganizmus, 3 - *Thermus thermophilus.*

Tairo Oshima, Department of Life Sciences, Tokyo Institute of Technology, Nagatsuta, Yokohama 226, Japan.

to the corresponding ones from mesophiles, though their stability to heat is unusually high. Table 1. illustrates examples of thermal stability of enzymes and tRNAs from *T. thermophilus* in comparison with those of mesophilic origin. Usually enzymes from *T. thermophilus* denatures at 90 °C or higher temperatures.

These heat stable enzymes from thermophiles are of industrial value and already some are used in diagnostic kits or as a component of commercial detergents or soups. Thermophile enzymes are also used as reagents in biochemical and molecular biological investigations. The best known example is DNA polymerase used in PCR. DNA polymerase from *T. thermophilus* is also sold under the name of Tth polymerase and is now widely used in DNA sequencing or *in vitro* DNA amplifications.

However it is more important to improve the stability of enzymes of industrial value than the direct use of thermophile enzymes. Some important enzymes can not be produced by thermophiles, and in this case artificial stabilization of the industrial enzyme is essential to develop the Enzyme Engineering. Enzymes used in industry can be stabilized based on knowledge obtained from studies on thermophile enzymes.

One example of such demand that improvement of heat stability of an industrial enzyme is highly required, is „glucose isomerase" used in the industrial production of fructose and syrup. Glucose isomerase used in the present industry is not stable over 60 °C. However, if we could succeed to improve the thermal stability of this enzyme and operate teh enzymatic process at higher temperatures, then we can expect the following three industrial merits. First of all, if the process is carried out at higher temperatures such as 80 °C, risk of microbial contamination during the isomerization process can be greatly reduced. The secondly, the equilibrium between glucose and fructose depends on the temperature and at higher temperature the production of fructose is more favored. Finally, viscosity is also depends on the temperature and viscosity decreases at higher temperatures. Thus troubles derived from highly viscsous solution of concentrated glucose is avoidable if the enzyme is stabilized.

Unfortunately our present knowledge on protein conformational stability is not sufficient to stabilize enzymes of industrial importance. many attempts have ben reported to stabilize proteins, however, there is no reliable theory how to stabilize protein. Our ability to design a robust enzyme is still immature and needs to improve. From these reasons, it is important to learn the molecular mechanisms by which thermophile enzymes are conferred their unusual stabilities against heat.

So far, many theories have been proposed to explain the molecular reasons why thermophile enzymes are stable to heat. But they can explain the stabilized mechanisms of only specified enzymesand proteins from thermophiles. No generalized or comprehensive theory exists. However, at moment, two principles are generally agreed among thermophile researches;

No 1 small changes are important for the increased stability of thermophile enzymes, and No 2 a variety of tactics can be applicable to improve thermal stability of an enzyme, and generally the effect of each tactic is additive, that is, if combined two different tactics, a more stabilized protein can be created.

We have focused our attention on the molecular mechanism of increased stability of enzymes from *T. thermophilus* and attempted to improve the stability of model enzymes based on our knowledge obtained from our analyses on the thermophile enzymes. In this article, two examples of stabilizing experiments will be described. In the first example, it will be shown how protein stabilization is difficult to understand and our knowledge and our knowledge on protein stability is not yet sufficient to understand the secrets of thermophile enzymes. In the second example, a new procedure to stabilze enzyme proteins using thermophiles willbe proposed which seems to be a promising way for protein designing and for understanding a protein structure-stability relationships.

Isopropylmalate dehydrogenase as a model protein

The experimental material chosen in the present study is isopropylmalate dehydrogenase (IPMDH). IPMDH is the third enzyme in leucine biosynthesis. The enzyme catalyzes hydrogen transfer to C-2 position of the substrate to NAD, and at the same time, the enzyme catalyzes the decarboxylation of COOH group attached to C-3 position of the substrate. The gene coding for this enzyme is usually called *leuB* gene. One reason why this enzyme was chosen is high sequence homology. So far *leuB* gene has cloned and sequenced from a variety of organisms including a weak, two moderate, two extreme, and a hyper thermophiles. Sequence homology is quite high and many homologous stretches are present throughout the sequence as shown in Fig.1. If deduced amino acid sequence of the *T. thermophilus* IPMDH is compared with that of a mesophile, for example, *Bacillus subtilis,* the sequence homology is more than 60 %. This high sequence homology made us possible to produce chimeric enzymes by *in vitro* crossing over of two *leuB* genes.

The gene for *T. thermophilus* enzyme was cloned, sequenced [5], and expressed in *E. coli* [6]. The enzyme produced in the mesophile was purified and enzymatic properties were extensively investigated [7]. The enzyme was also crrystallized and its three dimensional structure was solved at 1.9A resolution in cooperation with Prof. Katsube and his colleagues [8]. The enzyme is a homodimer and each subunit consists of 345 amino acid residues. Each subunit can be divided into two clearly isolated domains (see Fig.2.). The first domain contains both the N and C-terminals and is independent from the subunit-subunit interactions. The second domain consists of 150 residues and only this domain is involved in the subunit binding. A hydrophobic core is present at the center of the first domain, another hydrophobic core is at

Fig.1. Sequence comparison of IPMDHs. The first column represents standard abbreviation of microorganisms used for the comparative study.

Obr.1. Porovnanie sekvencií IPMDH. Prvý stĺpec predstavuje štandardné skratky pre mikroorganizmy používané v komparatívnej štúdii.

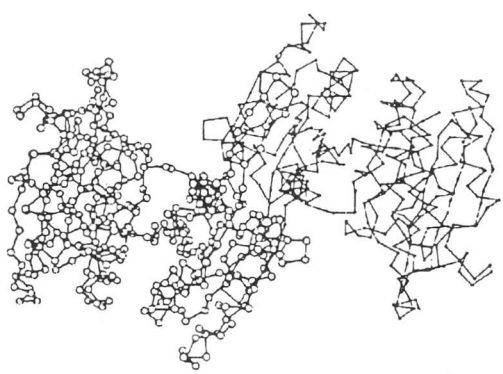

Fig.2. Backbone structure of *T. thermophilus* IPMDH.
Obr.2. Štruktúra hlavného reťazca *T. thermophilus* IPMDH.

the center of the second domain, and the third hydrophobic side chain clusterexist at the subunit/subunit interface.

Chimeric enzymes

Many chimeric IPMDHs have been constructed by connecting the thermophile and the mesophile, *B. subtilis leuB* genes [9]. Among chimeric enzymes, the most interesting one is 2T2M6T. The name designates the primary structure. The first 20 % from the N-terminal of chimera 2T2M6T is coded by the thermophile *leuB* gene, and the next 20 % is coded by the mesophile gene. The rest is again coded by the thermophile gene. The denaturation temperature of the chimeric enzyme is around 70 °C which is about 20 °C lower than that of the wild type thermophile enzyme. In the rest studies, this chimeric enzyme was used as the starting material, and attempts to stabilize this hybrid enzyme were made. In a sense, we have tried to recover this 20 °C difference in denaturation temperature by molecular design.

The chimera, 2T2M6T, was purified and crystallized. Three dimensional structure was solved with Prof. Katsube and Prof. Tanaka and their coleagues [10]. The backbone structure of the chimera is almost identical to that of the original thermophile enzyme except for two limited regions; one at around position 110. Both positions are in 2M part. Both positions are found in loop structure located outside of the molecule, that is, these two regions seem to be the most flexible parts of this enzyme.

Fig.3. shows the sequence of 2M part of the chimera in comparison with the corresponding sequence of the wild type enzyme of the thermophile. Though this part consists of 60 residues, due to high sequence homology only 22 residues are different between the chimera and the thermophile enzymes. If these 22 residues are replaced with the corresponding sequence of the thermophile IPMDH, thermal stability of the chimera can be improved to thelevel of the *T. thermophilus* enzyme. In a sense, 2T2M6T is a good exercise for

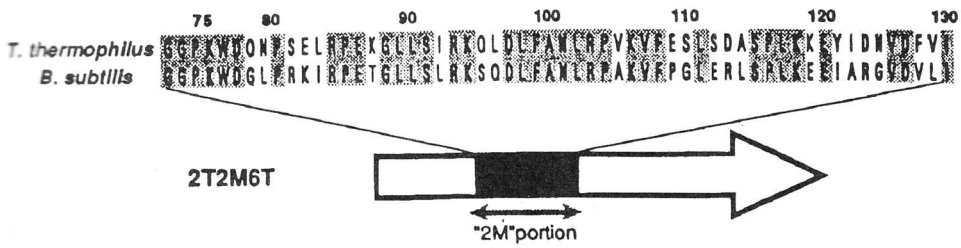

Fig.3. Amino acid sequence of 2M part.
Obr.3. Aminokyselinová sekvencia časti 2M.

protein chemists who are interested in molecular design of protein stabilization, because at least one answer is known.

Based on the three dimensional structure of the chimera, residue substitution at 82 was examined (Numataet al., manuscript in preparation). This position is occupied by Ser in case of 2T2M6T, but the same position is occupied by Arg in case of the extreme thermophile IPMDH. So a mutant S82R was made. Thermal denaturation of S82R was higher than that of chimera 2T2M6T. Thus a moleculr design based on crystallographic analysis seemed to be successful. The denaturation temperature was improved by 10 °C.

However, this successful mutant gave us a serious problem. When we locked into the side chain of position 82, it was turned out that this side chain is isolated from any other group. The position 82 is located on the surface of the enzyme, and does not interact with any part of the enzyme (Sakurai et al., in preparation). Arg-82 does not contribute to anyintramolecular interaction. We carefully analyzed the 3D structure in detail and fine structure showed that there is no bound water around Arg-82. So at moment the reason why Arg-82 is so important for the increased stability of the protein is unknown. This example indicates that conformational stability of a protein is still far from the comprehensive understanding, and theoretical design for a robust enzyme is still immature.

Module substitution

The substitution at position 110 was also investigated. Since remarkable displacement was found around position 110 when 3D structures of the chimera and the thermophile IPMDHs were compared as mentioned above. Residue-110 of the thermophile IPMDH is Pro, whereas the position is occupied by glutamate in the chimera. At first it was thought that this residue is extremelyimportant for the conformational stability of the protein. Position 110 is involved in a sharp turn. The turn consists of 4 residues from 110 to 113. Pro at 110 seems to fix the shape of this loop, and the shape is important to fix two beta strands: one up stream of this loop the other is down stream. And this fixation of two beta strands seemed to be crucial for the increased stability of the thermophile enzyme as well as the catalytic function.

Thus a mutant enzyme E110P is expected to be much more stable than the original chimera. The mutant seemed to be a quite reasonablemolecular design for improve the stability of the chimera. To tour surprise, mutant E110P did not show any improved stability (Numata et al., in preparation).

Careful analyses have been done concerning the conformational stability of the loop containing position 110. Position 111 is occupied by Gly in case of the thermophile IPMDH and torsion angle of Gly-111 is abnormal which is typical for Gly in beta turn structure. Thus it was considered that both Pro-110

and Gly-111 are important for the correct shape and also rigidity of this loop. Double mutant E110P/S111G was constructed by site directed mutagenesis, and our dissapointment was increased. Because the double mutant was more unstable than 2T2M6T is.

Finally it was noticed that Glu-113 forms a hydrogen bond with a backbone imino group. A triple mutant E110P/S111G/T113E was designed and constructed. The triple mutant drastically improved the stability of the chimera. In this case, any of single mutated protein did not show improved stability. Any of double mutation also did not improve the stability of 2T2M6T. Only triple mutant improved the thermal stability of chimera. The denaturation temperature was improved by about 10 °C. When added the first mutation, S82R, to this triple mutant, the resulted quadruple mutant showed almost same thermal denaturation profile as that of the original thermophile enzyme. In a sense, our original aim to recover the thermal stability of 2T2M6T was achieved by substituting four residues in loop regions.

The fact that only triple mutation improved the stability, but any single mutation did not improved the stability, indicates that there must be a structural unit in a protein. For stabilization, unit replacement seems to be more effectivethan single residue replacements. In this context, module analysis of a protin will be a promising way to improve the thermal stability. A module is defined as a compact unit of a protein consisting 10-40 continuous peptide fragment [11]. Module structure can be analyzed by measuring the distance between any two residues [12]. We analyzed module structure of IPMDH of *T. thermophilus* in cooperation with Prof. Go. The enzyme consists of 21 modules.

The loop from position 110 to 113 belongs to Module 7. By the triple mutation, sequence of Module 7 of the chimera coincide with that of the thermophile IPMDH except both ends. Thus the triple mutation can be regarded as module substitution. In short, it can be said that the chimera 2T2M6T was stabilized by replacement of module 7 rather than residue replacements. Module replacement would be more effective way than residue replacement for stabilizing a protein. Module analysis will give us a new clue for understanding the mechanism of increased stability of thermophile enzymes.

Evolutionary Molecular Engineering

Evolutionary Molecular engineering is a strategic concept to improve properties of biopolymers. The first step is to give random mutations to a desiredgene and the second step is to select and to reproduce a mutant gene under strong selective pressure, of which expression product acquired the desired property. The concept was first proposed by Eigen and Gardiner [13]. This method can be applicable for improving binding or catalytic properties as well as stability of an enzyme.

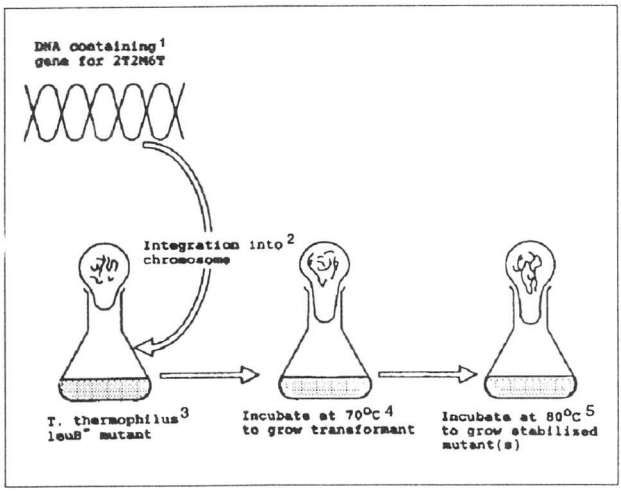

Fig.4. Experimental procedures to stabilized chimeric enzyme 2T2M6T in *T. thermophilus*.
Obr.4. Experimentálne postupy na stabilizáciu chimerického enzýmu 2T2M6T v *T. thermophilus*.
1 - DNA obsahujúca gén pre 2T2M6T, 2 - integrácia do chromozómu, 3 - *T. thermophilus leuB* mutant, 4 - inkubácia pri 70 °C pre rast transformantu, 5 - inkubácia pri 80 °C pre rast stabilizovaného mutantu (ov).

Chimera 2T2M6T was stabilized by evolutionary molecular engineering method (Tamakoshi, M. et al., unpublished data). The experimental procedures are schematically illustrated in Fig.4. The entire coding region of *leuB* gene was removed from chromosome of *T. thermophilus* strain HB27 by gene targeting method. If upstream and downstream sequences of *leuB* gene are connected to any foreign gene including one coding for chimera 2T2M6T, and the *leuB* minus mutant is treated with this DNA, the gene can be integrated in the position, where originally the coding frame of *leuB* gene is located, in leucine operon of the chromosomal DNA of the thermophile strain HB27. The integrated gene can be automatically expressed under the control of leucine operon promoter.

When 2T2M6T chimeric *leuB* gene was used as the foreign gene, the transformant was able to grow in the absence of leucine if the growth temperature was at or below 70 °C (the denaturation temperature of the chimera is around 70 °C) as expected (see Table 2.). After prolonged incubation at higher temperatures such as 76 or 79 °C, thermally stabilized mutants, that is, mutants capable of growing at the higher temperatures in the absence of leucine appeared as shown in Table 2.

3-Isopropylmalate dehydrogenases of the stabilized mutants were extracted and purified. Using the purified enzymes, it was confirmed that the gene products of these mutants are more stable than the original chimera, 2T2M6T. *leuB* genes of the mutants were sequenced. So far two mutated genes have been identified. Both are single base replacement which induces single amino acid replacement. One is from A to C in the chimeric *leuB* gene which replaces Ile93 (ATC) with Leu (CTC) in the mutated enzyme. In the thermophile enzyme, position 93 is ocupied with leucine, so that the laboratory evolution restored the thermophile sequence around this position. The three dimensional structure of the stabilized enzyme Ile93Leu has been studied

Table 2. Growth of transformant and mutants at various growth conditions.
Tabuľka 2. Rast transformantu a mutantov pri rôznych podmienkach rastu.

STRAINS [1]	+ Leucine [2]		- Leucine	
	70 °C	76 °C	70 °C	76 °C
HB27 (wild)	+	+	+	+
leuB minus [a]	+	+	-	-
Transformant [b]	+	+	+	-
Ile93Leu	+	+	+	+
Ala172Val	+	+	+	+

a - A *leuB* deleted mutant of strain HB27, b - The deletion mutant was transformed with the chimeric gene coding for 2T2M6T.
a - vynechaný mutant *leuB* kmeňa HB27, b - delečný mutant bol transformovaný chimerickým génom označeným ako 2T2M6T, 1 - kmene, 2 - leucín.

using X-ray crystallographic techniques (Tanaka, N. et al., manuscript in preparation). The study indicated that the backbone structure of the stabilized enzyme is almost identical with that of the original chimera 2T2M6T. Only difference detected is side chain conformation of the residue at position 93.

The other stabilized mutant contains thymine at position 515 instead of cytosine. As the result of this single base substitution, Ala-172 (GCG) is replaced with Val (GTG). The methyl group of Ala-172 resides in hydrophobic core of the second domain. Detailed inspection of the three dimensional structure of the thermophile 3-isopropylmalate dehydrogenase suggested that a small gap is present around the methyl group of Ala-172 and this gap seems to be filled by the mutation to Val.

These results obtained from the evolutionary molecular engineering studies strongly suggest that delicate arrangements such as correction of side chain conformation or filling up of small gap inside the hydrophobic core, are needed to stabilize a protein. Both mutations are not easy to design theoretically by simple inspection of the three dimensional structure. In a sense, the evolutionary molecularengineering method produced stabilized mutants which are hard to produce by theoretical methods.

References

1. BROCK, T.: Science, *230,* 1985, p. 132-138.
2. BROCK, T.: Thermophiles; General, Molecular, and Applied Microbiology, John Wiley & Sons, Inc., New York, 1986.
3. OSHIMA, T. - IMAHORI, K.: Intern. J. System. Bacteriol., *24 (1),* 1974, p. 102-112.
4. BORMAN, S.: Chem. Eng. News, Nov. 4, 1991, p. 31-34.
5. KIRINO, H. - AOKI, M. - AOSHIMA, M. - HAYASHI, Y. - OHBA, M. - YAMAGISHI, A. - WAKAGI, T. - OSHIMA, T.: Eur. J. Biochem., *220,* 1994, p. 275-281.

6. ISHIDA, M. - OSHIMA, T.: J. Bacteriol., *176 (9),* 1994, 2767-2770.
7. YAMADA, T. - AKUTSU, N. - MYIAZAKI, K. - KAKINUMA, K. - YOSHIDA, M. - OSHIMA, T.: J. Biochem., *108,* 1990, 449-456.
8. IMADA, K. - SATO, M. - TANAKA, N. - KATSUBE, Y. - MATSUURA, Y. - OSHIMA, T.: J. Mol. Biol., *222,* 1991, 725-738.
9. NUMATA, K. - MURO, M. - AKUTSU, N. - NOSOH, Y. - YAMAGISHI, A. - OSHIMA, T.:Protein Eng., *7 (12),* 1994.
10. ONODERA, K. - SAKURAI, M. - MORIYAMA, H. - TANAKA, N. - NUMATA, K. - OSHIMA, T. - SATO, M. - KATSUBE, Y.: Protein Eng., *7 (4),* 1994, 453-459.
11. GO, M.: Nature, *291,* 1981, 90.
12. NOGUCHI, T. - SAKAKIBARA, H. - GO, M.: Protein, *16,* 1993, 357.
13. EIGEN, M. - GARDINER, W.: Pure Appl. Chem., *53,* 1984, 967-978.

Do redakcie došlo 7.2.1995.

Stabilizácia enzýmových proteínov
s využitím poznatkov o termofilných mirkoorganizmoch a termofilných enzýmoch

TAIRO OSHIMA

Súhrn. Biopolyméry proteínov termofilného mikroorganizmu *Thermus thermophilus* HB 27 sa vyznačujú v porovnaní s biopolymérmi mezofilov (napr. *Bacillus subtilis*) vyššou termostabilitou. Potenciálne využitie termofilných enzýmov má pre priemyselné aplikácie mimoriadny význam. Na základe doterajších štúdií zákonitosti medzi štruktúrou a termostabilitou enzýmu 2 - izopropylmalátdehydrogenázy (2 - IPMD), využitím biotechnologických metód, modulovej analýzy a substitúcie sa stabilizoval chimérny enzým IPMP označený ako 2T2M6T. Proteosyntéza tohoto enzýmu klonovaným mezofilným kmeňom bola riadená génom, ktorý bol skonštruovaný spojením génov *leuB* z *Thermus thermophilus* a *leuB* z *Bacillus subtilis.* 2-IPMD expresovaná týmito génmi, je jedným z kľúčových pri biosyntéze leucínu.

Chimérny enzým 2T2M6T bol purifikovaný a vykryštalizovaný. Jeho nosná štruktúra bola takmer identická so štruktúrou pôvodného enzýmu 2-IPMP z *Thermus thermophilus* s výnimkou dvoch oblastí - v pozícii 82 a oblasti 110. Obe diferencie sú v časti štruktúry 2M, na slučke štruktúry molekuly proteínu. Študovala sa štruktúra vo vzťahu ku konformačnej stabilite slučky. Sekvenčnou analýzou sa identifikovalo poradie aminokyselín. Cielenou mutagenézou sa získali: mutant S82R, dvojnásobný mutant E 110P/S 111G. Až trojnásobná mutácia E 110P/S 111G/T 113E zvýšila termostabilitu chiméreho enzýmu o 10 °C. Adícia prvej mutácie k tejto trojnásobnej spôsobila, že termostabilita chiméreho enzýmu 2T2M6T produkovaného klovaným kmeňom bola rovnaká ako 2-IPMP produkovaného termofilom *Thermus thermophilus.* Pre stabilizáciu proteínu je teda potrebná zmena kompaktnej jednotky - modulu.

Metódou evolučného molekulového inžinierstva, na báze integrácie DNK obsahujúcu gén pre 2T2M6T s chromozomálnou DNK recipientného kmeňa *Thermus thermophilus*, sa získali mutanty schopné rásť pri teplotách 70 až 79 °C. Produkovaný enzým bol termostabilnejší než chiméra. Sekvenčnou analýzou kódujúceho jeho produkciu sa zistili zmeny v zložení aminokyselín. U jedného mutanta bol substituovaný izoleucín (I) leucínom (L) v polohe 93, u iného v pozícii 515 alanín (A) za valín (V). Analýza trojrozmernej štruktúry termofilnej 2-izopropylmalátdehydrogenázy röntgenovou analýzou preukázala prerušenie štruktúry v mieste okolo metylovej skupiny alanínu 172. Toto prerušenie bolo mutáciou za valín odstránené. Medzera sa nachádzala v hydrofóbnom jadre druhej domény molekuly proteínu.

Výsledky naznačujú, že citlivé zásahy do štruktúr proteínových biopolymérov vedú k ich stabilizácii.