Expression and Renaturation Analysis of Adenylosuccinate Synthetase from *Cordyceps cicadae*

Wenying Han and Haihua Ruan*

Tianjin Key Laboratory of Food Science and Biotechnology, College of Biotechnology and Food Science, Tianjin University of Commerce, Tianjin, China

Abstract. In this study, the objective was to obtain adenylosuccinate synthase (ADSS) in the adenosine biosynthesis pathway of *Cordyceps cicadae*. The coding sequence (CDS) information of the protein was obtained from transcriptome sequencing. Bioinformatics analysis were employed to predict and analyze the physicochemical properties, hydrophobicity, signal peptide, and advanced structure of ADSS. The ADSS coding sequence was chemically synthesized and a prokaryotic expression vector containing N-terminal His tag (pET28a-ADSS) was constructed. Induction of ADSS expression occurred with 0.1 mmol/L IPTG at 37°C for 2 hours. SDS-PAGE was utilized to detect soluble protein expression levels. Although ADSS primarily expressed as inclusion bodies, soluble proteins were successfully obtained after denaturation and renaturation.

1. Introduction

C. cicadae is a herb that is used in traditional Chinese medicine. *C. cicadae* is formed by the parasitization of *Paecilomyces cicadae* fungi on cicada nymphs, resulting in a complex of insects and fungi. This species mainly grows in Sichuan, Zhejiang, Jiangsu, Fujian, Yunnan, and other regions[1]. The processing method for *C. cicadae* has been documented as early as the Northern and Southern Dynasties[2]. *C cicadae* was dried in shade and sunlight, the soil was removed, and then boiled in pulp for a day, roasted and ground before use. *C. cicadae* contains active substances such as adenosine, polysaccharides[3], N^6-(2-hydroxyethyl)-adenosine (HEA), and ergosterol[4]. Among these compounds, adenosine is the primary active substance composed of ribose (furan ribose) and adenine moiety[5]. Adenosine participants have a variety of regulatory processes *in vivo*, with functions in regulating sleep[6], targeting tumors[7], anti-cancer[8], anti-inflammatory[9], improving cardio-cerebral blood circulation[10, 11], and inhibiting the nervous system[12].

Compared with chemical synthesis, synthetic or semi-synthetic bioactive substances synthesized by *in vitro* enzymatic reaction are more suitable for large-scale, high yield and green production. We used inosine monophosphate (IMP) and aspartic acid (ASP) to generate Adenosuccinate under the action of ADSS, and then Adenosine monophosphate (AMP) as generated under the action of ADSL. The pathway of adenosine production by AMP under the action of NT5E constructs the enzyme reaction system of adenosine in vitro.ADSS catalyzes the conversion of IMP to adenylosuccinate and is one of the key enzymes in the adenosine biosynthesis pathway. However, there are few reports on ADSS genes in *C. cicadae*.

In this study, the CDS of ADSS was obtained through transcriptome sequencing. Based on this, various online analysis tools were utilized to predict the physicochemical properties and structural domains of the gene and its encoded protein. The expression vector pET28a was employed, along with *Escherichia coli* BL21 (DE3), for heterologous expression of ADSS from *C. cicadae*. Soluble proteins were acquired by extracting, washing, denaturing, dissolving, and renaturing the inclusion bodies formed by these proteins. This provides a reliable protein refolding technology for obtaining soluble proteins.

2. Materials and methods

2.1. Sequence properties of ADSS in *C. cicadae*

ADSS gene information was retrieved from transcriptome data. The comparison results on the NCBI website showed that the sort by percent identity between the results and adenylosuccinate synthetase (*Isaria fumosorosea* ARSEF 2679) was 82%, and its gene ID was 30022817. The ExPasy Protoparam tool (http://web.expasy.org/protparam/) was used to calculate sequence properties including the amino acid sequence length, molecular weight, isoelectric point and grand average of hydropathicity (GRAVY) of ADSS protein. The subcellular localization, signal peptides, transmembrane helices, secondary and tertiary structure of ADSS were predicted using PSORT (https://www.genscript.com/psort.html), SignalP (http://www.cbs.dtu. dk/services/SignalP/), DeepTMHMM (https://dtu.biolib.com/DeepTMHMM), SOPMA (https://npsaprabi.ibcp.fr/cgibin/npsaautomat.pl?page=npsa_sopma.html) and SWISS-MOD-

* Corresponding author: ruanhaihua@tjcu.edu.cn

EL (https://swissmodel.expasy.org/interactive), respectively. Qingke Biotechnology Co., Ltd. was commissioned to optimize the CDS of ADSS based on codon bias of *E. coli*. The recombinant His_6-ADSS protein was constructed by chemical synthesis and connected to the BamHI and XhoI sites of pET-28a expression vector.

2.2. Protein expression, solubilization, and refolding

E. coli were cultivated at 37℃ and 220 rpm until reaching an optical density (OD_{600nm}) of approximately 0.6-0.8, followed by induction of expression using 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for a duration of 2 h at the same temperature and rotational speed. Subsequently, the cells were harvested by centrifugation at 5000×g and kept at a temperature of 4℃. To resuspend the cells, they were mixed with a Tris-HCl buffer pH8.0 containing 150 mM NaCl, 1 mM protease inhibitor phenylmethylsulfonyl fluoride (PMSF). Ultrasonication was employed to disrupt *E. coli* membranes and release cellular contents into the solution. Following ultrasonication, separation between cell debris and protein extract was achieved through centrifugation for duration of ten minutes at a speed of 12000×g while maintaining the temperature at 4℃. The resulting sediment and supernatant fractions were subjected to electrophoresis under denaturing conditions utilizing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Prior to loading onto the gel matrix, samples were mixed with treatment buffer composed of 250mM Tris-HCl pH6.8, 50% Glycerine, 0.2% Bromophenol Blue, 10% SDS, and 5% 2-mercaptoethanol, and subsequently heated for ten minutes at 100℃. Finally, the gels were stained with Coomassie Brilliant Blue R-250 dye to visualize the proteins. The inclusion bodies containing His_6-ADSS were obtained by centrifuging the total cell lysate at 12000×g for 10 minutes at 4℃. To solubilize these inclusion bodies, an 8M urea solution was employed. Subsequently, the denatured His_6-ADSS proteins in this solution underwent refolding and dialysis following previously described protocols[13].

3. Results and discussion

3.1. Construction of ADSS prokaryotic expression vector

Different kinds of microorganisms have different codon usage bias, and codon optimization can further improve the expression level of recombinant protein in *E. coli*. After optimization, the CDS codon adaptation index of ADSS was increased from 0.40 to 0.85, and then the optimized fragment was chemically synthesized and linked to the pET-28a vector between BamHI and XhoI sites (Fig. 1).

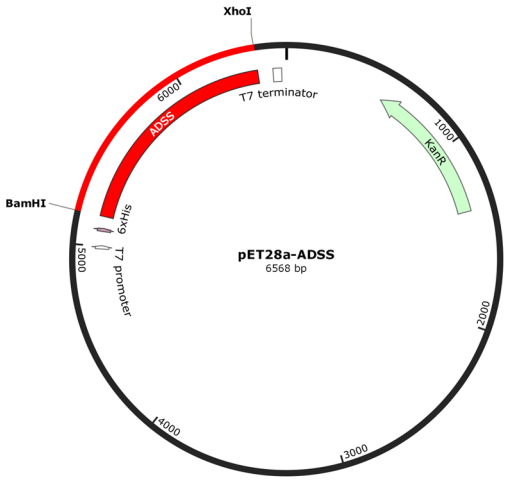


Fig. 1. Plasmid map of pET28a-ADSS

3.2. Physicochemical properties analysis of ADSS

Protein physicochemical properties of ADSS are shown in Table 1, ADSS consists of 410 amino acids with a theoretical molecular mass of 44.48 kDa and a pI value of 7.72. In terms of hydrophilicity and hydrophobicity, the maximum value for glycine at position 115 and the minimum value for aspartic acid at position 52 in the ADSS protein polypeptide chain are 2.311 and -3.078, respectively. The overall average hydrophilicity and hydrophobicity value is -0.129. As shown in Fig.2, there are more hydrophilic amino acids with negative values throughout the polypeptide chain, indicating that the ADSS protein polypeptide chain is predominantly hydrophilic in terms of its hydrophilicity and hydrophobicity properties. The analysis revealed a probability of zero for the presence of a signal peptide in the ADSS protein, indicating that it is a non-secretory protein without a signal peptide sequence.

Table 1. Protein physicochemical characteristics of ADSS.

Protein name	ADSS
Size (aa)	410
MW (kDa)	44.48
PI	7.72
GRAVY	-0.129
Signal peptide	no

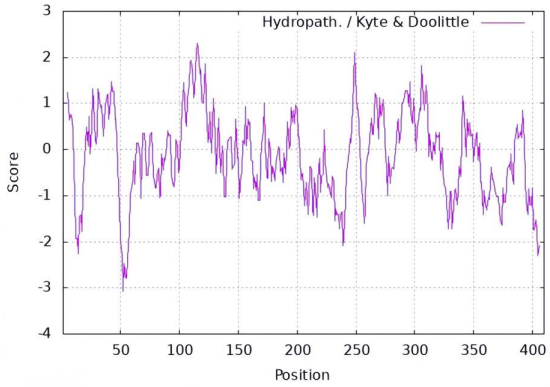


Fig. 2. Hydrophilicity and hydrophobicity prediction of ADSS

3.3. Protein transmembrane structure of ADSS

The transmembrane structure of proteins can be analyzed to determine the protein's location, which significantly impacts subsequent expression and purification processes. The DeepTMHMM online analysis software was used to predict the transmembrane domain of ADSS proteins, and the results are presented in Table 2. The number of predicted transmembrane helices (TMH) is 0. The probability that the N-terminus is located on the cytoplasmic side or inside the cell membrane is 0. The expected number (Exp) of amino acids (AAs) in TMHs of ADSS protein was calculated as 8.62 (less than 18), and there were no transmembrane amino acids within the first 60 amino acids of the protein, indicating an absence of transmembrane regions in these proteins.

Table 2. Transmembrane domain prediction of ADSS.

TMH	Value
Number of predicted TMH	0
Exp number of AAs in TMHs	8.62
Exp number, first 60 AAs	0
Total prob of N-in	0.38
Outside	1

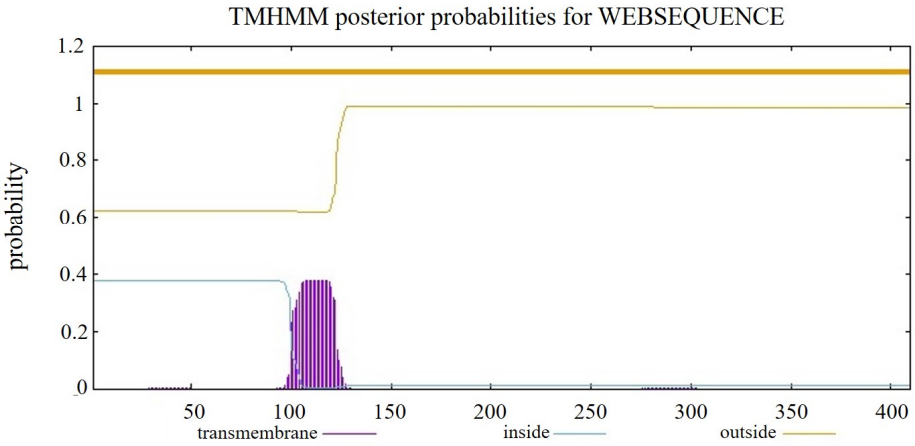
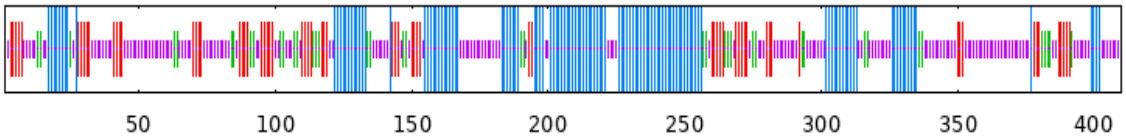


Fig. 3. Transmembrane domain prediction of ADSS.



Blue is used to represent α-helix, green for β-turn, yellow for random coil, red for extended strand, and the numbers below indicate the positions of the corresponding amino acids.
Fig. 4. Secondary structure prediction of ADSS.

3.4. Prediction of secondary and tertiary structure of ADSS

The secondary structure prediction revealed that the protein consists of 30.49% α-helix, 9.76% β-turn, 16.59% extended strand, and 43.17% random coils (Table 3 and Fig.4). Additionally, the online SWISS MODEL website was utilized to perform homology modeling for tertiary structure prediction (Fig.5). Notably, the tertiary structure of ADSS protein is predominantly composed of random coil and α-helix elements, which aligns with the results obtained from secondary structure prediction analysis.

Table 3. The distribution ratio of protein secondary structures

Secondary structures	ADSS
α-helix	30.49%
β-turn	9.76%
Random coils	43.17%
Extended strand	16.59%

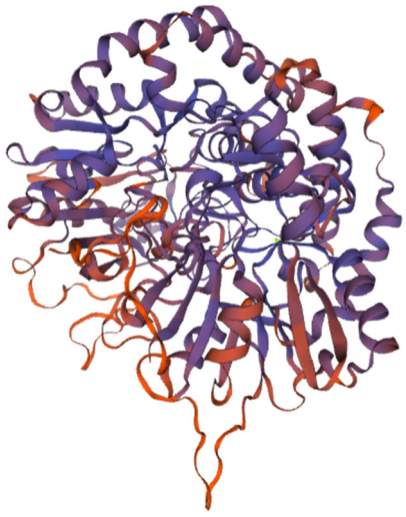


Fig. 5. Tertiary structure prediction of ADSS

3.5. Subcellular localization

The subcellular localization of a protein can predict its location within the cell, which is highly significant for protein research. The predicted intracellular localization of the ADSS protein by PSORT is presented in Table 4 below. The results indicate that ADSS is predominantly located in the cytoplasm and mitochondria, with a high probability, suggesting potential processing modifications involving the Golgi apparatus may be required.

Table 4. Subcellular localization

Subcellular location	ADSS
Cytoplasmic (%)	34.8
Plasma membrane (%)	-
Mitochondrion (%)	30.4
Golgi Body (%)	13.0
Endoplasmic Reticulum (%)	8.7
Nucleus (%)	4.3
Extracellular (%)	4.3
Vesicles of secretory system (%)	4.3

3.6. Preparation of standard curves for protein concentrations

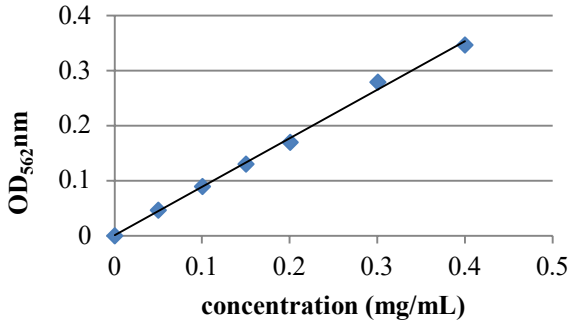


Fig. 6. Standard curve of protein concentration.

The BCA protein concentration detection kit was utilized for determining the protein content. The standard curve relating BSA concentration to absorbance value was obtained:

$$y = 0.8821x + 0.0008$$

$$R^2 = 0.9968$$

The y represents the absorbance value and x denotes the protein concentration (mg/mL).

3.7. SDS-PAGE analysis of ADSS proteins

Lane M serves as the standard protein marker, with the molecular masses (kDa) indicated on the left side of the gel. Lanes 1 and 2 correspond to the total cell lysates of *E. coli* expressing pET28a-ADSS induced with 0 and 0.1 mM IPTG, respectively. Lanes 3 and 4 represent the supernatant and precipitate obtained from ultrasonication at 4℃ followed by centrifugation at 12000×g for 10 min at 4℃ of the *E. coli* cell lysate.

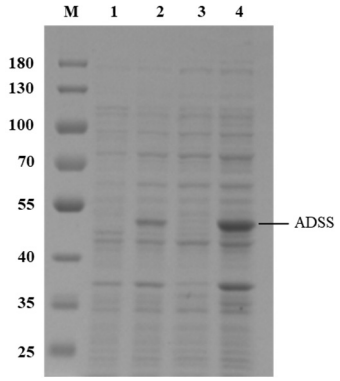
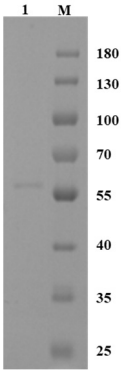


Fig. 7. SDS-PAGE analysis of ADSS protein expression

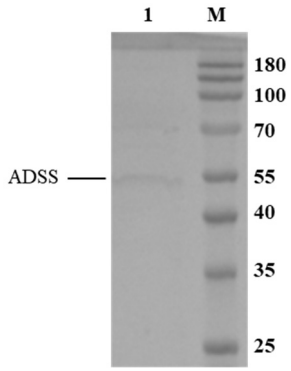
To investigate the activity of the ADSS protein, we generated a fusion construct comprising an N-terminal His_6 tag (His_6-ADSS) that encompassed the open reading frame (ORF) of the ADSS protein. The His_6-ADSS was successfully expressed in E. coli cells and exhibited a migration pattern corresponding to a 44.48kDa protein on 10% SDS-PAGE, consistent with its predicted amino acid sequence (Table 1). However, it was observed that the expressed His_6-ADSS predominantly existed in an insoluble form as no soluble fraction could be detected in the supernatant of cells expressing His_6-ADSS (compare lane 3 in Fig.7). This observation suggests that the expressed His_6-ADSS likely forms inclusion bodies within bacterial cells. It is common for recombinant proteins produced in transformed microorganisms to form insoluble and inactive inclusion bodies. In our case, Subcellular localization predictions indicated that ADSS may require glycosylation and protein modification on the Golgi apparatus, which is absent in the *E.coli* expression system (Table 4). In the prediction of transmembrane domains, it is possible that a transmembrane domain may exist between the 100th and 150th amino acids(Fig.3). Incorrect protein folding can result in the exposure of the hydrophobic transmembrane domain to the exterior, which could subsequently lead to the formation of inclusion bodies. Consequently, fast expression rates coupled with incorrect folding environments led to insoluble inclusion body formation.

3.8. Renaturation of inclusion bodies



Lane M is standard protein markers, and the molecular masses (kDa) are indicated to the right of the gel. Lane 1 represents the His_6-ADSS protein after refolding.

Fig. 8. SDS-PAGE analysis of ADSS protein after renaturation



Lane M is standard protein markers, and the molecular masses (kDa) are indicated to the right of the gel. Lane 1 represents the His_6-ADSS protein after dialysis.

Fig. 9. SDS-PAGE analysis of ADSS protein after dialysis

The solubilization of inclusion bodies was conducted using 8M urea, following the protocol previously described [12]. We initially disrupted bacterial cells expressing His_6-ADSS by ultrasonication in 50mM Tris-HCl (pH8.0) supplemented with 150mM NaCl and 1mM PMSF at a temperature of 4℃. Subsequently, the inclusion bodies containing His_6-ADSS were obtained as a pellet by centrifuging the total *E. coli* cell lysate at 12000×g for 10 minutes at 4℃. To dissolve the inclusion bodies present in the pellet, we performed a prewash using cold Tris-HCl solution containing 2M urea, 1mM EDTA, and 0.5% Triton X-100 along with a mixture of 0.1M NaCl and 10mM dithiothreitol (DTT). Following this step, we successfully solubilized the inclusion bodies in cold Tris-HCl solution comprising 8M urea, 150mM NaCl and 20mM DTT. Consequently, nearly all His_6-ADSS proteins present in the inclusion bodies were effectively dissolved. To facilitate the refolding of His_6-ADSS proteins, the aforementioned solution containing denatured pET28a-ADSS proteins was gradually diluted by adding a refolding buffer (composed of Tris-HCl with 5% glycerol, 500mM NaCl, 1mM DTT and 0.5mM EDTA) until the final urea concentration reached 0.5 M. Subsequently, the mixture was incubated on ice for an additional duration of 24h. To eliminate the denaturant urea from the system, dialysis was performed on the refolded protein solution using a volume of 1 L 50 mM Tris-HCl, 5% glycerol, and 500mM NaCl at a temperature of 4℃ over a period of 48 h with four hanges of dialysis buffer.

The limitation of utilizing inclusion body proteins lies in the absence of effective refolding techniques, with emphasis placed on protein rearrangement and refolding. In this study, the reverse dilution method was initially employed to remove urea denaturant, followed by addition of renaturation buffer to the denaturation solution. This simultaneous reduction in denaturant and protein concentration facilitated gradual rearrangement of intermediate protein structures[14]. After renaturation on ice for 24h, the denaturant concentration was gradually and uniformly reduced by redialysis. This method avoids that when the concentration of denaturant drops to a low level, part of the misfolded and aggregated proteins cannot be structurally rearranged to their native state, and produces a local high concentration of denatured proteins, which is more conducive to inhibiting aggregation.

The ADSS proteins were denatured, renatured (Line 1 in Fig.8), and then dialyzed (Line 1 in Fig.9) to form soluble forms. The ADSS protein exhibited a distinct band at approximately 44kDa as determined by SDS-PAGE, with no other stray bands observed. The protein concentration of ADSS measured by standard curve of protein concentration after dialysis was 0.88mg/mL (Fig.6).

4. Conclusion

In this study, bioinformatics analysis of ADSS was conducted, and a prokaryotic expression plasmid containing the open reading frame of ADSS was successfully constructed and highly expressed in *E. coli* BL21(DE3). Simultaneously, high purity soluble ADSS protein was obtained through denaturation and renaturation. Subsequently, we will optimize the refolding method to obtain a more active form of the ADSS protein and demonstrate its enzymatic activity by catalyzing the reaction between IMP and ASP [15]. The foundation was laid for constructing in vitro enzymatic reaction to produce adenosine.

References

1. F. Di, J. Zhang, W. Cheng, et al. *Paecilomyces cicadae*;: a systematic overview of the biological activities and potential mechanisms of its active metabolites [J]. Food and Agricultural Immunology, **34**.(2023)
2. K. Liu, F. Wang, G. Liu, et al. Effect of Environmental Conditions on Synnema Formation and Nucleoside Production in Cicada Flower, *Isaria cicadae* (Ascomycetes) [J]. Int J Med Mushrooms, **21**: 59-69.(2019)
3. T. Zizhong, L. Wenjie, C. Yusheng, et al. Extraction, Purification, Physicochemical Properties, and Activity of a New Polysaccharide From *Cordyceps cicadae* [J]. Frontiers in nutrition, **9**: 911310-911310.(2022)
4. S. Fenhui, L. He, J. Qian, et al. Optimization of the nutritional constituents for ergosterol peroxide production by *Paecilomyces cicadae* based on the uniform design and mathematical model [J]. Sci Rep, **12**: 5853.(2022)
5. O. J. Olatunji, Y. Feng, O. O. Olatunji, et al. Neuroprotective effects of adenosine isolated from *Cordyceps cicadae* against oxidative and ER stress damages induced by glutamate in PC12 cells [J]. Environmental Toxicology and Pharmacology, **44**: 53-61.(2016)
6. L. Michael, C. Jiang-Fan, H. Zhi-Li, et al. Adenosine and Sleep [J]. Handbook of experimental pharmacology, **253**: 359-381.(2019)
7. A. Ryan, L. Robert, N. Aung, et al. Next steps for clinical translation of adenosine pathway inhibition in cancer immunotherapy [J]. Journal for

immunotherapy of cancer, **10**.(2022)

8. M. Shuli, L. Yingying, Y. Lijuan, et al. Potential and promising anticancer drugs from adenosine and its analogs [J]. Drug discovery today, **26**: 1490-1500.(2021)
9. P. Silvia, C. Chiara, B. P. Andrea, et al. Adenosine and Inflammation: Here, There and Everywhere [J]. International Journal of Molecular Sciences, **22**: 7685-7685.(2021)
10. K. Heena, K. Parneet, S. T. Gurejet, et al. Adenosine as a Key Mediator of Neuronal Survival in Cerebral Ischemic Injury [J]. Neurochemical research, **47**: 3543-3555.(2022)
11. R. Allison, G. David, K. Lora, et al. Adenosine and the Cardiovascular System [J]. American journal of cardiovascular drugs : drugs, devices, and other interventions, **19**: 449-464.(2019)
12. Zhou, Mengmeng, Wu, et al. Adenosine signaling mediate pain transmission in the central nervous system [J]. Purinergic Signalling, **19**: 1-10.(2022)
13. H. Ruan, Z. Yan, H. Sun, et al. The YCR079w gene confers a rapamycin-resistant function and encodes the sixth type 2C protein phosphatase in *Saccharomyces cerevisiae* [J]. FEMS yeast research, **7**: 209-215.(2007)
14. S. M. Singh, A. K. Panda. Solubilization and refolding of bacterial inclusion body proteins [J]. J Biosci Bioeng, **99**: 303-310.(2005)
15. A. Bubić, M. Narczyk, A. Petek, et al. The pursuit of new alternative ways to eradicate *Helicobacter pylori* continues: Detailed characterization of interactions in the adenylosuccinate synthetase active site [J]. Int J Biol Macromol, **226**: 37-50.(2023)