CLONING AND EXPRESSION OF RECOMBINANT THYMIDINE PHOSPHORYLASE IN *PICHIA PASTORIS*

Jaloliddin Abdurakhmanov*, Sobirdjan Sasmakov, Shukhrat Khasanov, Oybek Ashirov, Farkhod Eshboev, Khayotjon Dolimov, Saodat Gaynazarova, Sadullaev Tulkin, Yarilkaganova Aygul, Shakhnoz Azimova

S.Yu. Yunusov Institute of the Chemistry of Plant Substances, Academy of Sciences of the Republic of Uzbekistan, Mirzo Ulugbek Str. 77, 100170, Tashkent, Uzbekistan

* Correspondence: jaloliddin0919@mail.ru; Tel.: +998 97157 9944;

Abstract

Based on the transfer vector pPICZαA the recombinant plasmid DNA pPICZαA-TP (4862 bp) containing cDNA (deoA gene, 1326 bp) of the Thymidine phosphorylase (TP) of *E. coli* were constructed. The cloned plasmid pPICZαA-TP used for expression of recombinant thymidine phosphorylase in *Pichia pastoris* yeast. The substrate specificity of the corresponding recombinant enzyme was studied by hydrolysis of thymidine (2-Deoxy-β-D-ribofuranosyl) thymine) to thymine and 2-deoxy-β-D-ribofuranosyl. The recombinant thymidine phosphorylase (≈49 kD) expressed in *Pichia pastoris* GS115 cells exhibits high hydrolytic activity and can be used for enzymatic transglycosylation of nucleosides.

Keywords: *nucleoside phosphorylases, modified nucleosides, recombinant plasmid, Pichia pastoris, thymidine phosphorylase.*

Introduction

Modified nucleosides are heterocyclic nitrogenous bases of natural or synthetic origin containing monosaccharides - cyclic pentoses. Modified nucleosides (or nucleoside analogues) are used for the treatment of viral diseases and certain forms of cancer (Molnupiravir - SARS-CoV-2, Tenofovir - HIV, HBV, Sofosbuvir - HCV, Fludara, Cladribine - leukaemia of various aetiologies etc.). Nucleoside analogues can be synthesized by chemical or enzymatic methods, or a combination of these methods [1-4]. Chemical synthesis is a long multi-stage process, including the introduction and removal of various protective groups in the carbohydrate residue and heterocyclic base, which leads to a significant decrease in the efficiency of the process [2,5]. At present, in the world, many drugs based on modified nucleosides are obtained by methods of multi-stage chemical synthesis, which has a number of significant disadvantages. Enzymatic synthesis has a number of advantages over chemical synthesis: mild reaction conditions, high stereoselectivity, minimal use of polluting chemicals and organic solvents, high efficiency, and the absence of unwanted impurities. Therefore, the biotechnological method for the synthesis of nucleosides using genetically engineered nucleoside phosphorylases makes it

possible to replace chemical synthesis with enzymatic [6-7]. The enzymatic transglycosylation reaction is usually carried out at 50°C to inhibit other enzymes such as deaminases. At this temperature, some nucleoside phosphorylases retain most of their activity. For example, reactions involving thymidine phosphorylase are usually carried out at 45°C [2].

Bacterial glycosyltransferases catalyzing the transfer of a pentofuranose group to purine or pyrimidine bases are successfully used in the synthesis of various natural nucleoside analogues of biological and pharmaceutical importance [2,8]. The use of microorganisms producing nucleoside phosphorylases for the synthesis of modified nucleosides of biological and pharmaceutical importance has proved to be highly effective. An analysis of the literature has shown that *Escherichia coli* thymidine phosphorylase (*Ec*TP, EC 2.4.2.4) is currently successfully used for the biocatalytic synthesis of derivatives of heterocyclic nitrogen-containing bases [9]. At present, one of the advanced expression systems that allow to obtain recombinant proteins on an industrial scale is the yeast system of *Pichia pastoris*. The accumulation of significant biomass during cultivation on inexpensive nutrient media, the absence of endotoxins and pyrogens, a higher level of recombinant protein synthesis compared to other expression systems are the advantages of *Pichia pastoris* yeast. [10-13]. Based on this, the aim of this work is to clone a recombinant plasmid encoding *Escherichia coli* thymidine phosphorylase (*Ec*TP) in the *Pichia pastoris* expression system.

Materials and methods

In the experiments, enzymes and reagents from New England Bi-olabs (USA), Thermo Fisher Scientific (USA), Sigma-Aldrich (USA), SBS Genetech (China), Biosset (Russia) and Himedia (India) were used.

Isolation of genomic DNA of *Escherichia coli*

To obtain genomic DNA samples, *Escherichia coli* strain RKMUz – 221 (the strain was obtained from the collection of industrial microorganisms of the Institute of Microbiology of the Academy of Sciences of the Republic of Uzbekistan) cells were grown in liquid nutrient medium (LB Broth Miller, Invitrogen) at 37°C for 12 h. Isolation was carried out in a 1.5 ml tube containing 100 µl of bacteria in a liquid nutrient medium. Centrifuged at 8500 rpm for 4 min. 90 µl of the supernatant was removed and the bacteria were suspended in the remaining 10 µl. Add 90 µl PEG-NaOH reagent solution (60 ml PEG 200, 5 ml 2 M NaOH and 25 ml dH2O) [14]. The mixture was heated at 80°C for 10 min and

shaken briefly before use. One-microliter samples were used directly for PCR in 20 µl reactions.

Sample preparation for ligation

Thymidine phosphorylase cDNA (deoA gene, insert), size of 1326 bp was amplified by PCR using as a template genomic DNA isolated from cells of *Escherichia coli* strain RKMUz - 221 using the following primers:

1) Forward – 5'- AATTGAATTCATGTTGTTTCTCGCACAA -3'

2) Reverse – 5'- TTTTTCTAGATTATTCGCTGATACGG -3'

The amplificon (insert) was purified by precipitation with 0.5 M magnesium chloride (in final concentration 0.05 M) and 96% ethanol (in final concentration 70%) [15].

One µg of the transfer vector pPICZαA and PCR amplificates were sequentially treated with FastDigest *EcoRI* and FastDigest *XbaI* restrictases (Thermo Scientific, USA). Linear vector DNA with a length of 3536 bp isolated after electrophoresis in 0.8% agarose gel using PureLink Quick Gel Extraction Kit (Invitrogen) [16].

Vector ligation with insert (thymidine phosphorylase gene)

Ligation of the pPICZαA vector with the target gene - thymidine phosphorylase was carried out in a volume of 10 µl in a molar ratio of 1:3, respectively, using recombinant T4 DNA ligase (Invitrogen).

Transformation of *E. coli* TOP10 cells with pPICZαA-TP ligand

After precipitation with the ligation mixture, the pellet (recombinant DNA) was dissolved in 10 µl of de-ionized water. Dissolved DNA was transformed into electrocompetent *E. coli* cells of the Top 10 plasmid-free strain (Invitrogen) according to the method [17].

Identification of recombinant clones

The molecular weight of plasmid DNA isolated from clones obtained after transformation with a ligation mixture was determined by electrophoresis in 0.8% agarose gel. Clones containing DNA with a length of 4862 bp (deoA gene + pPICZαA transfer vector fragment) were inspected by PCR analysis. PCR amplicons were analyzed by 1% agarose gel electrophoresis. [18].

Transformation vector (pPICZαA-TP) and expression of recombinant protein in *Pichia pastoris* GS115

The recombinant pPICZαA-TP vector was linearized with *SacI* and transformed into *Pichia pastoris* GS115 using a Electroporator 2510 (Eppendorf, USA). The yeast was then grown in YPD selection medium containing the antibiotic Zeocin™ (100 µg/mL). To induce

thymidine phosphorylase expression, *Pichia pastoris* GS115 strains were inoculated in BMGY medium at 30°C with vigorous shaking (220 rpm), and then transferred to BMMY and incubated at 30°C with vigorous shaking (220 rpm) for 72 hours. Methanol (0.5-1%, v/v) was supplied to the culture medium once a day [21].

Purification of recombinant protein, SDS-PAGE

Extraction of the enzyme was carried out by destroying the yeast cells using a benchtop two-stage high pressure homogenizer (APV - 2000, SPX Flow, USA). The device is capable of treating liquid samples at pressures up to 2000 bar. For the treatment of yeast suspensions, the two-stage (two times 1200 bar) was used. The suspension temperature at the exit of the homogenizer did not exceed 40°C. Treated samples were cooled to 15°C immediately at the outlet of the device using a water-ice bath. After that, the yeast culture liquid medium was added to the homogenate. To obtain primary purified thymidine phosphorylase homogenate was precipitated using gradient precipitation with 20%-60% ammonium sulphate. The precipitate containing the enzyme was separated by centrifugation at 8000 rpm for 30 min. The precipitated enzyme was dissolved in PBS (150mM NaCl, 5.2mM Na_2HPO_4, 1.7mM KH_2PO_4), pH 7.4.

The molecular weight was determined by electrophoresis in 10% PAGE in the presence of Na dodecyl sulphate (SDS) according to the Laemmle method [19]. The protein concentration was determined by the Lowry method, using bovine serum albumin (BSA) as the standard [20].

Enzymatic hydrolysis of Thymidine by recombinant TP

Reaction mixture contained 3.3 mM KH_2PO_4, pH 8.0, 2 mM thymidine (1-(2-Deoxy-β-D-ribofuranosyl) thymine) (Sigma-Aldrich, Cat. T9250-25G) and 5 µl thymidine phosphorylase fraction (total protein cons. 16.7 mg/ml). The reaction mixtures were incubated at 50°C up to 8 hours. Substrate and product quantities were determined using HPLC.

HPLC detection

The products of hydrolysis by the thymidine phosphorylase were detected using High Performance Liquid Chromatography (HPLC): Shimadzu LC-20A Prominence, column Shim-pack GIST C18 5 µm, 4.6 × 150 mm, eluents A – 10 mM KH_2PO_4 in H2O, B - MeCN, flow rate 0.5 mL/min, detection PDA at 254 nm. Gradient: 0 min - 100% A, 10 min -95% A, 15 min – 90% A, 20 min - 100% A.

Results and discussion

Recombinant plasmid DNA pPICZαA-TP (4862 bp) was constructed based on the transfer vector pPICZαA (3593 bp) containing a 1326 bp thymidine phosphorylase (deoA) gene encoding *Escherichia coli* thymidine phosphorylase (49 kDa by PAGE electrophoresis) for expression in the yeast *Pichia pastoris*.

Figure 1 shows the genetic map of the cloned plasmid DNA pPICZαA-TP, which includes the following elements: 5'AOX1 promoter (methanol-regulated induction promoter): nucleotides location 2-940 (939 bp); α-factor gene for the secretion of the expressed protein into the nutrient medium: nucleotides location 941 - 1207 (267 bp); deoA gene (cDNA encoding *E. coli* thymidine phosphorylase): nucleotides location 1214–2539 (1326 bp); *Myc* epitope tag (human *c-Myc* proto-oncogene, intended for affinity purification of the expressed protein): nucleotides location 2544 - 2573 (30 bp); *HIS* tag (polyhistidine tag for affinity purification of the expressed protein): nucleotides location 2589 - 2606 (18 bp); transcription termination fragment 3'AOX1 TT: nucleotides location 2686 - 2932 (247 bp); BleoR gene of resistance to zeocin for selection of recombinant clones: nucleotides location 3432 - 3806 (375 bp) [21].

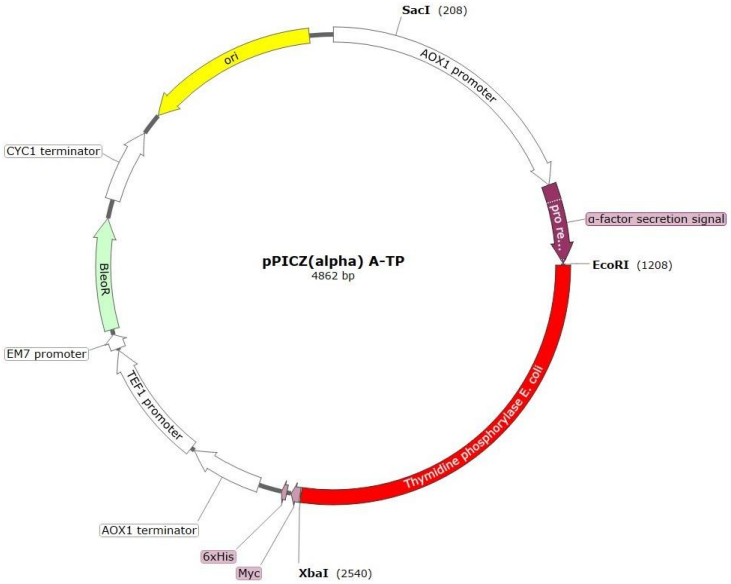


Figure 1. Physical map of the constructed recombinant plasmid pPICZαA-TP

The obtained clones were screened by PCR using primers synthesized for gene amplification. Figure 2 shows the result of PCR analysis of recombinant clones. According to the results of PCR, the selected clones contain the DNA fragment (gene) with a molecular weight of 1326 bp (**Fig. 2, A**). Thus, PCR indicated the presence of an insert (the deoA gene) in the investigated plasmid.

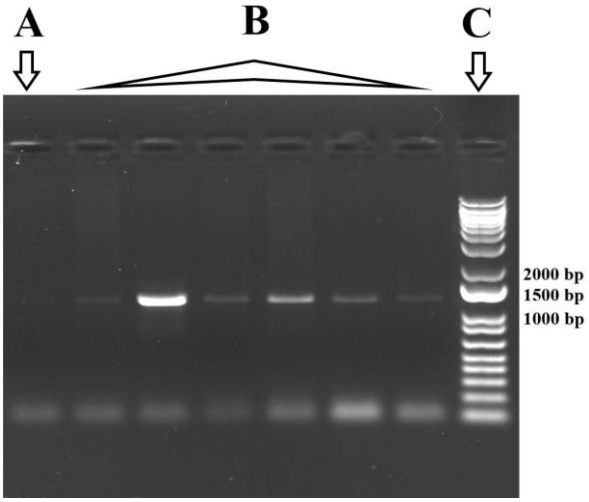


Figure 2. PCR analysis of recombinant clones

A – pPICZαA (K-); **B** – pPICZαA- TP; **C** – DNA Marker.

Figure 3 shows a comparative restriction analysis of the recombinant plasmid pPICZαA-TP and the original pPICZαA vector at a single *SacI* restriction site. After electrophoresis analysis, linear DNA were detected, the mass of which corresponds to theoretical calculations according to the physical maps of plasmids.

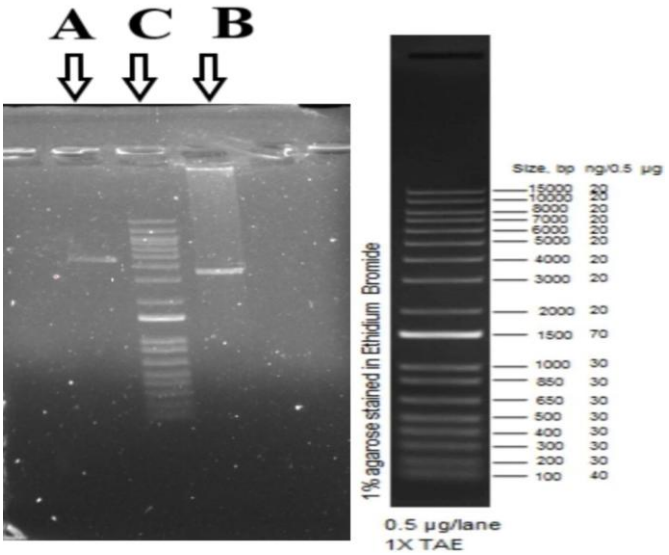


Figure 3. Restriction analysis of the recombinant plasmid pPICZαA-TP

A – pPICZαA-TP/SacI (4862 bp); **B** – pPICZαA/SacI (3593 bp); **C** – DNA Marker.

Expression of the recombinant enzyme was carried out in 250 ml four baffled flasks for 72 hours. After destroying the yeast cells by high pressure homogenizer, homogenate was cooled to 15°C immediately in a water-ice bath. The yeast culture liquid medium was added to the homogenate. Homogenate was separated by centrifuging at 8000 rpm for 30 min after gradient precipitation with 20%-60% ammonium sulphate and the fraction were dissolved in PBS pH 7.4. The fraction containing recombinant *E. coli* thymidine phosphorylase from the

yeast *Pichia pastoris* GS115 using denaturing 10% SDS-PAGE electrophoresis was analyzed by staining with Coomassie G-250 (**Figure 4**). The protein concentration in the fraction containing the recombinant enzyme, determined by the Lowry method, is 16.7 mg/ml. Figure 4 shows that the recombinant enzyme appeared in the 49 kDa band corresponding to its nucleotide sequence.

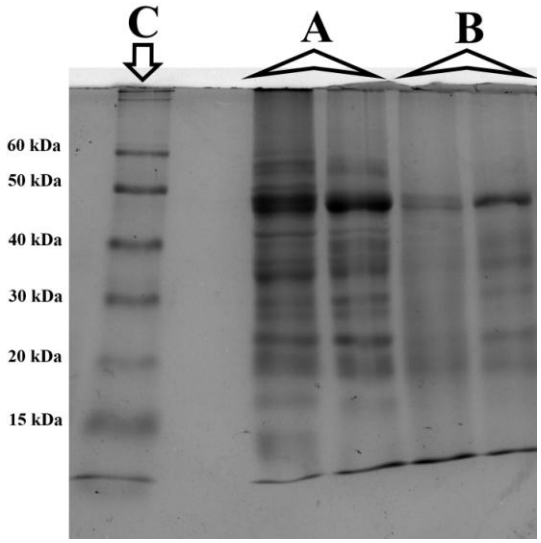
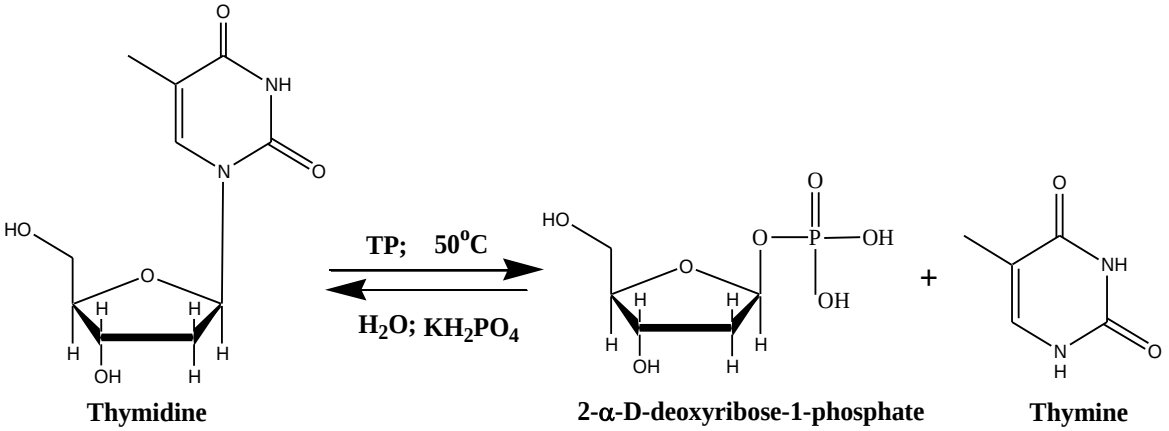


Figure 4. 60% Ammonium sulphate fraction of *Pichia pastoris* protein extract

A – *Pichia pastoris* protein extract (60 and 40 µg protein per well); **B** – fraction after gradient precipitation with 20%-60% ammonium sulphate (3 and 7 µg protein per well); **C** – Novex Sharp Pre-Stained Protein Standard (Invitrogen).

The catalytic activity of the fractions contained thymidine phosphorylase was determined by hydrolysis of thymidine (**Scheme 1**).



Scheme 1. Enzymatic hydrolysis of Thymidine (2-Deoxy-β-D-ribofuranosyl) thymine) by thymidine phosphorylase (TP)

The level of hydrolysis and the identification of the formed thymine after hydrolysis were carried out by HPLC for 1-6 hours (supplementary file, figure 5).

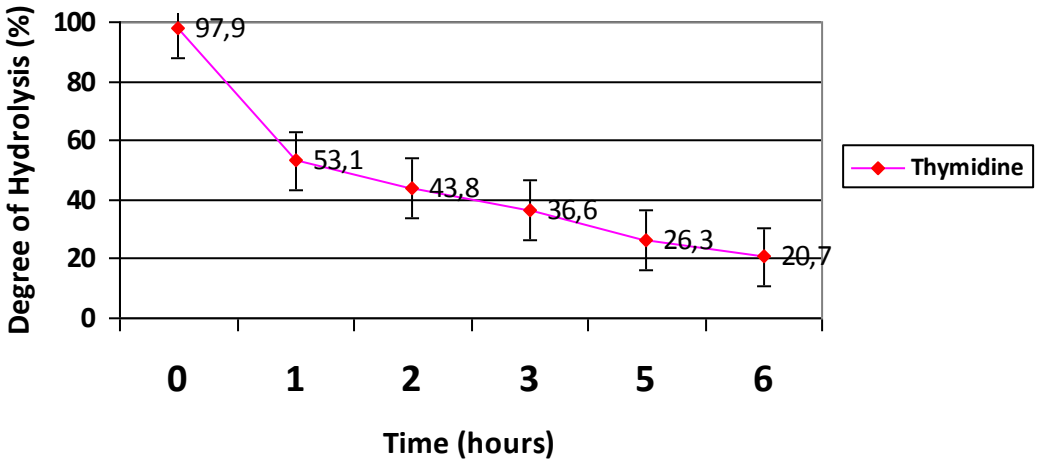


Figure 5. Time-course of enzymatic hydrolysis of Thymidine catalysed by TP

As can be seen from the graph (Fig. 5), the fraction containing recombinant TP leads to the hydrolysis of thymidine within 1 hours by more than 50% under our experimental conditions, which confirms its high hydrolytic activity.

Conclusions

Thus, a new recombinant plasmid DNA pPICZαA-TP (4862 bp) for expression of recombinant thymidine phosphorylase in the yeast *Pichia pastoris* was cloned. The achieved results indicate that thymidine phosphorylase (≈49 kD) has substrate specificity to thymidine (2-Deoxy-β-D-ribofuranosyl) thymine) and may be used for enzymatic transglycosylation reactions of thymidine (pyrimidine) type nucleosides.

Acknowledgements

The study was carried out within the framework of project F-FA-2021-360, Ministry of Innovative Development of the Republic of Uzbekistan.

Conflict of interest

The authors declare no conflicts of interest.

References

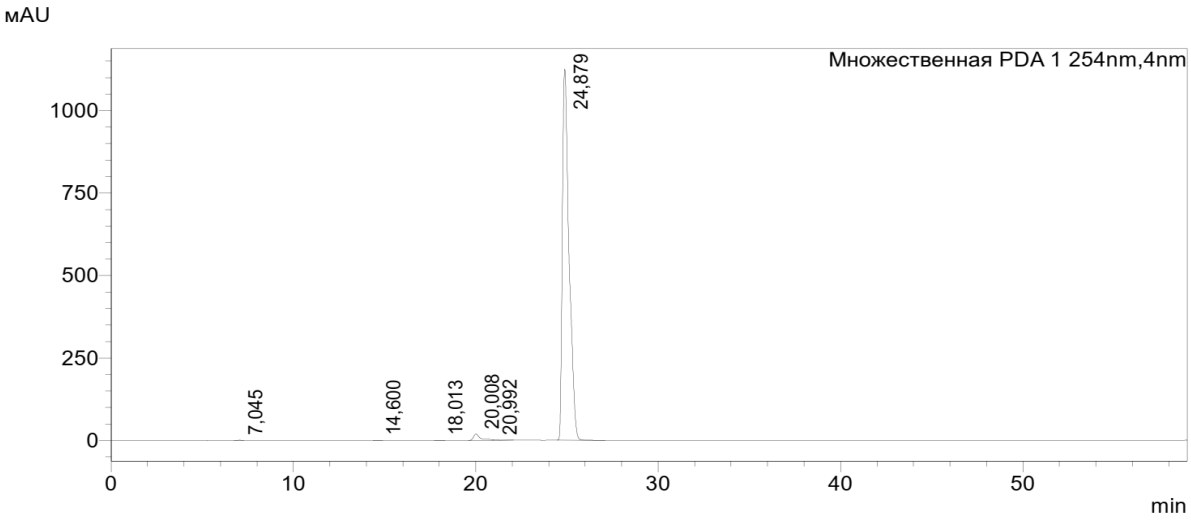
1. National Center for Biotechnology Information. PubChem Compound Summary for CID 657237, Fludarabine. (2022). Retrieved from https://pubchem.ncbi.nlm.nih.gov/compound/Fludarabine.
2. Fateev, I.V., Kostromina, M.A., Abramchik, Y.A., Eletskaya, B.Z., Mikheeva et al. (2021) Multi- Enzymatic Cascades in the Synthesis of Modified Nucleosides: Comparison of the Thermophilic and Mesophilic Pathways. *Biomolecules* **11**, 586. doi.org/10.3390/biom11040586

3. Cox, R.M., Wolf, J.D., Plemper, R.K. (2021) Therapeutically administered ribonucleoside analogue MK-4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. *Nat. Microbiol* **6**, 11-18. doi.org/10.1038/s41564-020-00835-2
4. Nucleotide polymerase inhibitors Sofosbuvir // Essential Medicines WHO Model List. (2015) - 19th ed. - Geneva: WHO. Retrieved from http://www.who.int/medicines/publications/essentialmedicines/EML2015_8-May-15.pdf
5. Fernández-Lucas, J., Fresco-Taboada, A., Acebal, C., de la Mata, I., Arroyo, M. (2011) Enzymatic synthesis of nucleoside analogues using immobilized 20-deoxyribosyltransferase from Lactobacillus reuteri. *Appl. Microbiol. Biotechol*. **91**, 317-327.
6. Fernández-Lucas, J. (2015) Multienzymatic synthesis of nucleic acid derivatives: A general perspective. *Appl. Microbiol. Biotechol*. **99**, 4615-4627.
7. Mikhailopulo, I.A., Miroshnikov, A.I. (2010) New Trends in Nucleoside Biotechnology. *Acta naturae* **2**(5), 36-58.
8. Del Arco, J., Fernández-Lucas J. (2018) Purine and pyrimidine salvage pathway in thermophiles: A valuable source of biocatalysts for the industrial production of nucleic acid derivatives. *Appl. Microbiol. Biotechnol*. **102**, 7805-7820.
9. Kamel, S., Yehia, H., Neubauer, P., Wagner, A. (2019) *Enzymatic Synthesis of Nucleoside Analogues by Nucleoside Phosphorylases* (First Edition) Wiley-VCH Verlag GmbH & Co., Germany, pp. 1-28.
10. Sasmakov, S.A., Ashirov, O.N., Abdurakhmanov, J.M., Khasanov, Sh.Sh., et al. (2021) Expression of recombinant PreS2-S protein from the hepatitis B virus surface antigen in Pichia pastoris. *VacciMonitor* **30**(1), 27-32.
11. Ahmad, M., Hirz, M., Pichler, H., Schwab, H. (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol*. **98**, 5301-5317.
12. Spohner, S.C., Quitmann, H., Czermak, P. (2015) Expression of enzymes for the usage in food and feed industry with *Pichia pastoris*. *Journal of Biotechnology*. **202**, 118.
13. Gasser, B, Prielhofer, R, Marx, H, Maurer, M, Nocon, J, Steiger, M, Puxbaum, V, Sauer, M, Mattanovich, D. (2013) *Pichia pastoris*: protein production host and model organism for biomedical research. *Future Microbiol*. **8**(2), 191-208. doi.org/10.2217/fmb.12.133.

14. Laurie, V., Eric, H. F. (2015) A single protocol for extraction of gDNA from bacteria and yeast. *BioTechniques* **58**, 120-125 doi.org/10.2144/000114263
15. Shupeng, He, Bozhi Cao, Yi Yi, Shenhao, Huang et all. (2022) DNA precipitation revisited: A quantitative analysis. **Nano select 3**(3), 617-626. doi.org/10.1002/nano.202100152
16. PureLink™ Quick Gel Extraction Kit. *Molecular cloning*. Catalog number: K210012.
17. Sambrook, Joseph. (2001). *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, pp. 1.119
18. Sambrook, Joseph. (2001). *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, pp. 5.4
19. He, F. (2011) Laemmli-SDS-PAGE. *Bio-Protocol* **1**, e80.
20. Waterborg, J.H., Matthews, H.R. (1984) *Methods in Molecular Biolog, Proteins: The Lowry Method for Protein Quantitation, vol 1* Humana Press. doi.org/10.1385/0-89603-062-8:1
21. EasySelect™ Pichia Expression Kit. (2010) User Manual. For Expression of Recombinant Proteins Using pPICZ and pPICZα in *Pichia pastoris*. Manual part no. 25-0172.

Supplementary file

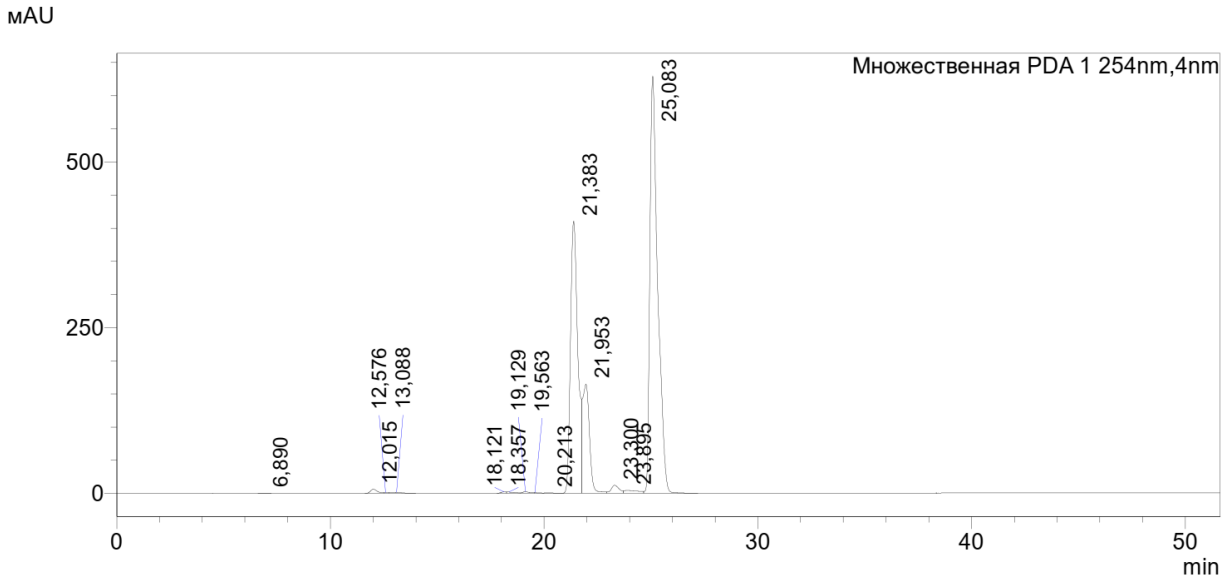
1) Hydrolysis of Thymidine – 0 min.



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	7,045	11862	961		0,040
2	14,600	6021	390		0,021
3	18,013	3795	213		0,013
4	20,008	526912	19327		1,799
5	20,992	43015	1796		0,147
6	24,879	28704963	1123771	Thymidine	97,981
Сумма		29296569	1146457		100,000

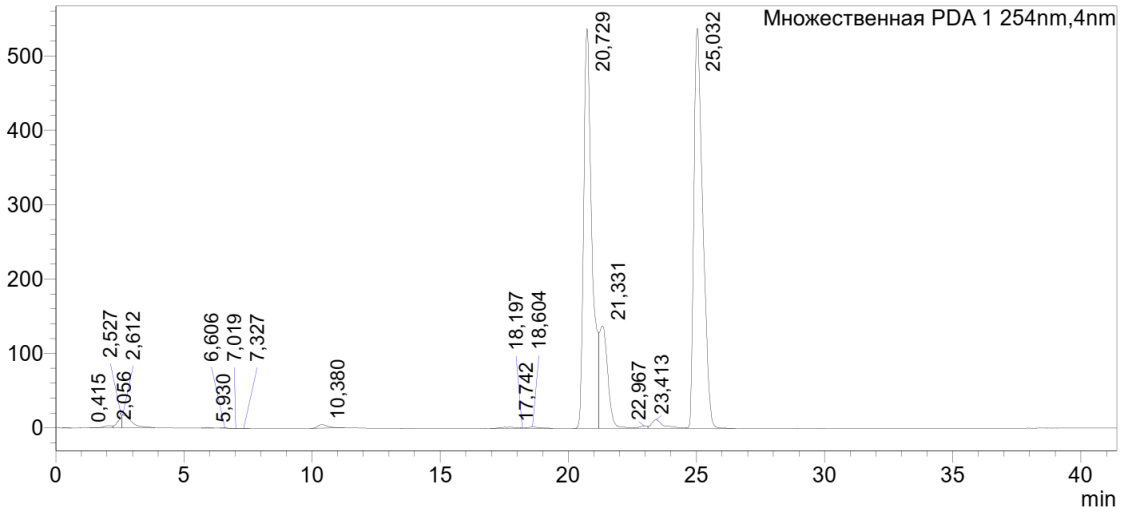
2) Hydrolysis of Thymidine – 1 h.



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	6,890	4364	258		0,014
2	12,015	133644	6139		0,438
3	12,576	15619	1047		0,051
4	13,088	45560	1259		0,149
5	18,121	26898	1707		0,088
6	18,357	41276	1560		0,135
7	19,129	49934	2140		0,164
8	19,563	10738	751		0,035
9	20,213	11304	408		0,037
10	21,383	9768931	410242	Thymine	32,042
11	21,953	3709333	164377		12,167
12	23,300	306848	11938		1,006
13	23,895	195121	4113		0,640
14	25,083	16168068	628568	Thymidine	53,032
Сумма		30487637	1234506		100,000

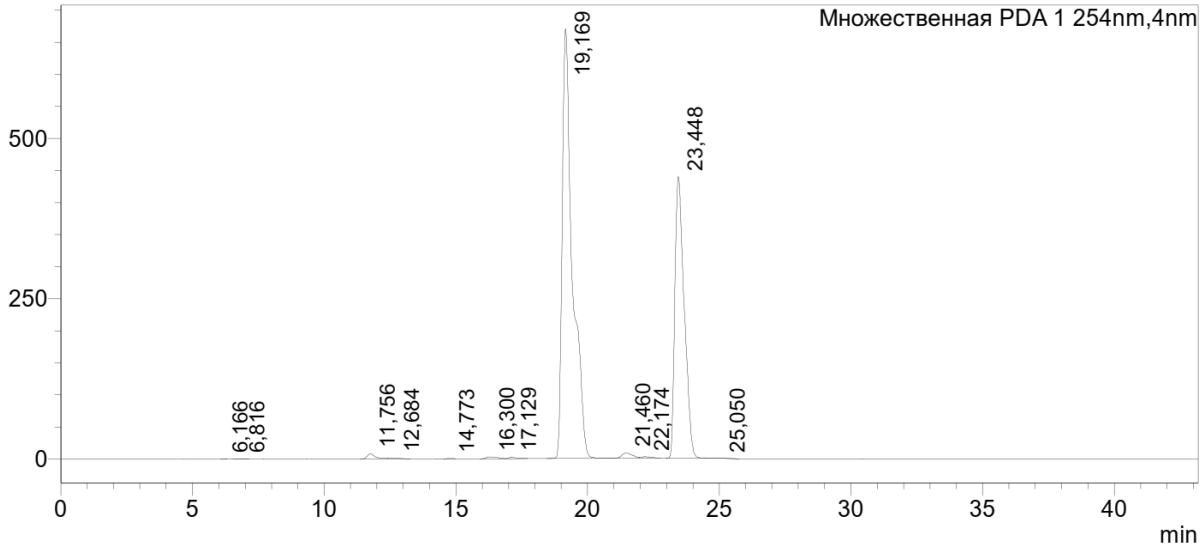
3) Hydrolysis of Thymidine – 2 h.



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	0,415	3213	394		0,011
2	2,056	62119	2726		0,209
3	2,527	197698	19925		0,667
4	2,612	409562	17159		1,381
5	5,930	3601	279		0,012
6	6,606	7649	787		0,026
7	7,019	1673	109		0,006
8	7,327	1454	135		0,005
9	10,380	131345	4692		0,443
10	17,742	69563	1510		0,235
11	18,197	10881	912		0,037
12	18,604	59037	2065		0,199
13	20,729	12096402	536981	Thymine	40,781
14	21,331	3154375	137681		10,634
15	22,967	68369	3072		0,230
16	23,413	378309	11548		1,275
17	25,032	13006692	537428	Thymidine	43,850
Сумма		29661940	1277404		100,000

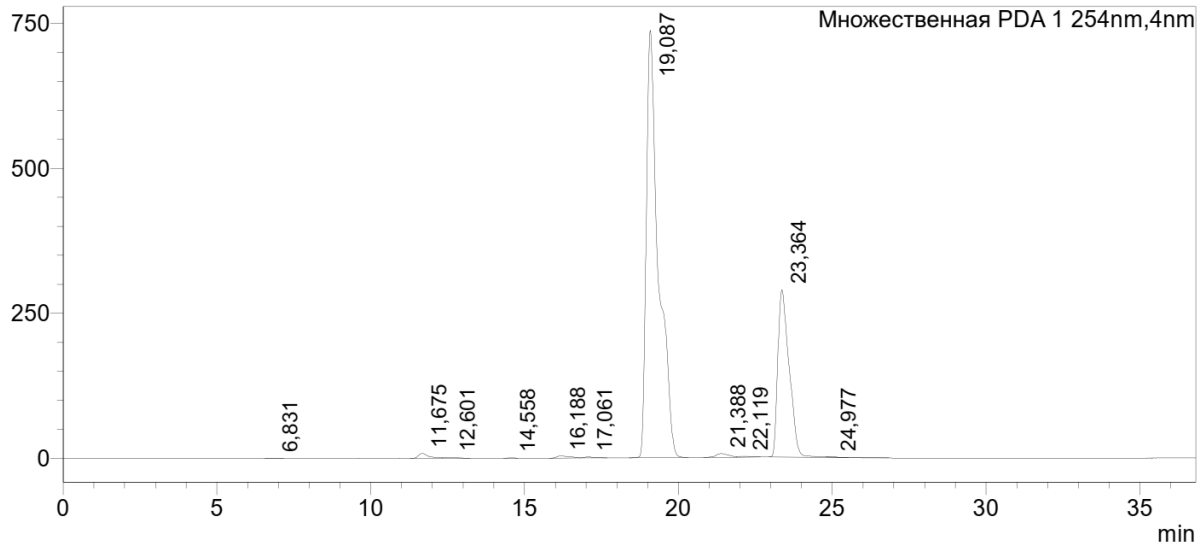
4) Hydrolysis of Thymidine – 3 h.



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	6,166	1029	122		0,003
2	6,816	4383	253		0,015
3	11,756	185536	7916		0,627
4	12,684	48238	1798		0,163
5	14,773	9376	684		0,032
6	16,300	89900	3086		0,304
7	17,129	49384	1968		0,167
8	19,169	18054757	669535	Thymine	61,059
9	21,460	228014	8207		0,771
10	22,174	51884	2052		0,175
11	23,448	10834861	438777	Thymidine	36,642
12	25,050	12010	498		0,041
Сумма		29569371	1134896		100,000

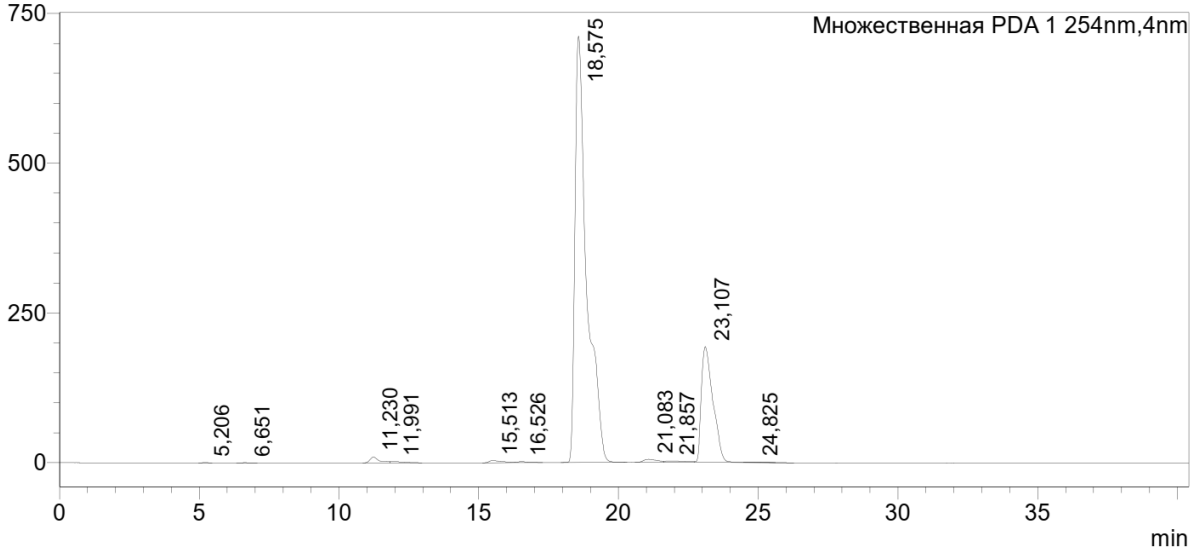
5) **Hydrolysis of Thymidine – 5 h.**



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	6,831	4185	239		0,015
2	11,675	204404	8554		0,722
3	12,601	57529	2004		0,203
4	14,558	8461	622		0,030
5	16,188	113328	3811		0,400
6	17,061	46459	1782		0,164
7	19,087	20216020	736418	Thymine	71,395
8	21,388	170384	6187		0,602
9	22,119	37867	1536		0,134
10	23,364	7448343	288378	Thymidine	26,305
11	24,977	8855	391		0,031
Сумма		28315835	1049923		100,000

6) **Hydrolysis of Thymidine – 6 h.**



PDA Ch1 254nm

Peak#	Ret. Time	Area	Height	Name	Площадь%
1	5,206	2495	169		0,009
2	6,651	6821	346		0,025
3	11,230	224362	9813		0,823
4	11,991	69485	2072		0,255
5	15,513	120435	3650		0,442
6	16,526	47092	1701		0,173
7	18,575	20848997	711133	Thymine	76,508
8	21,083	159110	4778		0,584
9	21,857	109452	2283		0,402
10	23,107	5645484	193211	Thymidine	20,717
11	24,825	16919	533		0,062
Сумма		27250652	929688		100,000