

Article

# Comparative Analysis of Mesophilic YqfB-Type Amidohydrolases

Roberta Statkevičiūtė<sup>1,\*</sup>, Mikas Sadauskas<sup>1</sup> , Juta Rainytė<sup>2</sup> , Karolina Kavaliauskaitė<sup>1</sup>  
and Rolandas Meškys<sup>1,\*</sup> 

<sup>1</sup> Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Sauletekio av. 7, 10257 Vilnius, Lithuania

<sup>2</sup> Department of Eukaryote Gene Engineering, Institute of Biotechnology, Life Sciences Center, Vilnius University, Sauletekio av. 7, 10257 Vilnius, Lithuania

\* Correspondence: roberta.statkeviciute@gmc.vu.lt (R.S.); rolandas.meskys@bchi.vu.lt (R.M.)

**Abstract:** The widespread superfamily of the human activating signal cointegrator homology (ASCH) domain was identified almost 20 years ago; however, the amount of experimental data regarding the biological function of the domain is scarce. With this study, we aimed to determine the putative cellular functions of four hypothetical ASCH domain-containing amidohydrolase YqfB analogues by investigating their activity towards various *N*-acylated cytosine derivatives, including potential nucleoside-derived prodrugs, as well as their ability to bind/degrade nucleic acids in vitro. According to determined kinetic parameters, *N*<sup>4</sup>-acetylcytidine is assumed to be the primary substrate for amidohydrolases. Despite the similarity to the proteins containing the PUA domain, no nucleic acid binding activity was detected for YqfB-like proteins, suggesting that, in vivo, these enzymes are a part of the pyrimidine salvage pathway. We also demonstrate the possibility of the expression of YqfB-type amidohydrolases in both prokaryotic and eukaryotic hosts. The small protein size and remarkable halotolerance of YqfB-type amidohydrolases are of great interest for further fundamental research and biotechnological applications



**Citation:** Statkevičiūtė, R.; Sadauskas, M.; Rainytė, J.; Kavaliauskaitė, K.; Meškys, R. Comparative Analysis of Mesophilic YqfB-Type Amidohydrolases. *Biomolecules* **2022**, *12*, 1492. <https://doi.org/10.3390/biom12101492>

Academic Editor: Judit Toth

Received: 6 September 2022

Accepted: 13 October 2022

Published: 16 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** ASCH; amidohydrolase; *N*<sup>4</sup>-acetylcytidine; prodrug

## 1. Introduction

The human activating signal cointegrator-1 homology (ASCH) domain is widespread in all domains of life and some prokaryotic viruses [1]. It has been identified mainly at the C-terminus of ASC-1 (also known as thyroid receptor-interacting protein 4 (TRIP4)), which is associated with nuclear receptors and transcription factors and is part of an RNA-interacting complex [2]. ASCH superfamily proteins consist of ~110 amino acids and contain a conservative octapeptide G-X-K-X-X-[ETS]-X-R. ASCH domain-containing proteins have been shown to adopt a tertiary fold, which was originally observed in the RNA-binding pseudouridine synthase and archeosine transglycosylase (PUA) protein superfamily, even with low protein sequence similarity. Based on the known properties and functions of homologous domain-containing proteins in RNA metabolism, the ASCH domain is thought to be able to interact with RNA molecules as a transcription co-regulator [3,4]. However, there is a lack of available experimental data that highlight the biochemical characterization and biological functions of this diverse protein superfamily.

In a previous study, a hypothetical ASCH domain-containing protein from *Zygomonas mobilis* was described as a monomeric ribonuclease that binds nucleic acids and degrades single-stranded RNA molecules in vitro [5,6]. Recently, the *Escherichia coli* ASCH domain-containing protein YqfB was characterized as the smallest known amidohydrolase with only 103 amino acids [7]. The primary substrate of the enzyme is the modified nucleoside *N*<sup>4</sup>-acetylcytidine (ac4C), which is abundant in both eukaryotic and prokaryotic RNA molecules. The presence of ac4C is important for efficient protein synthesis and

eukaryotic ribosome biogenesis [8–12]. The cytidine acetylation writer enzymes have been well studied in both prokaryotes and eukaryotes, but the prokaryotic erasers have not yet been identified.

Enzymes belonging to the large class of amidohydrolases are known to hydrolyze various linear and cyclic amides. These biocatalysts are attractive in biotechnology for their potential application in chemoenzymatic synthesis of pharmaceuticals. The amidohydrolase YqfB was shown to hydrolyze a well-known cancer prodrug, capecitabine, as well as  $N^4$ -acylated 5-fluorocytosines (cytidines), suggesting potential application in a selective two-step activation system for the prodrug-targeted therapy.

Here we describe ASCH domain-containing proteins from mesophilic bacteria *Buttiauxella agrestis* (BagYqfB; 109 a.a.), *Cronobacter universalis* (CunYqfB; 113 a.a.), *Klebsiella pneumoniae* (KpnYqfB; 103 a.a.), and *Shewanella loihica* (SloYqfB; 107 a.a.). A comparative study showed that the recombinant proteins, including *E. coli* YqfB, were highly salt tolerant, and active towards a broad substrate range of  $N^4$ -acylated cytosines and cytidines. In addition, we present data on the production of a functional recombinant YqfB from *E. coli* in the yeast *Pichia pastoris*.

## 2. Materials and Methods

### 2.1. Phylogenetic Analysis of ASCH Domain-Containing Proteins

In total, 1082 ASCH-domain-containing protein sequences were retrieved from the non-redundant protein sequences (nr) database using BLASTp search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>, accessed on 6 October 2022). The initial query was based on the YqfB amino acid sequence. The BLASTp search algorithm was set to exclude the *Escherichia* and *Klebsiella* genera, and to filter sequences outside the 50–80% sequence identity and 95–100% query coverage ranges. Multiple sequence alignment of these sequences, including the sequences of the YqfB and KpnYqfB, was performed with Clustal Omega [13] under default settings. The phylogenetic tree was visualized with iTOL v5 [14]

### 2.2. Strains and Chemicals

Cell cultures of the target microorganisms *Buttiauxella agrestis* DSM 9389, *Cronobacter universalis* DSM 27963, and *Shewanella loihica* DSM 17748 were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). *Klebsiella pneumoniae* strain KV-3 is a veterinary isolate [15].

The *E. coli* DH5 $\alpha$  (Novagen, Germany) strain was used for routine cloning experiments, while for the recombinant synthesis of target proteins, *E. coli* HMS174  $\Delta$ pyrF  $\Delta$ yqfB cells were used. PichiaPink<sup>TM</sup> expression system (Invitrogen, Carlsbad, CA, USA) strain 2, which is an *ade2* and *pep4* knockout, was used for YqfB expression in *Pichia pastoris*.

Cytidine,  $N^4$ -acetylcytosine,  $N^4$ -acetylcytidine,  $N^4$ -acetyl-2'-deoxycytidine, and  $N^4$ -isobutyryl-2'-deoxycytidine were purchased from Combi-Blocks (San Diego, CA, USA).  $N^4$ -benzoylcytidine, capecitabine, 5'-fluorocytosine, and *p*-nitrophenyl acetate were obtained from Sigma Aldrich (St. Louis, MO, USA). Cytosine was purchased from Alfa Aesar (Germany), and isocytosine, 2'-deoxycytidine, and 5'-fluorocytidine were from Biosynth Carbosynth (Berkshire, UK).

The preparation of  $N^2$ -acetylisocytosine [7] and  $N^4$ -hexanoyl-2'-deoxycytidine,  $N^4$ -nicotinoyl-2'-deoxycytidine,  $N^4$ -(3-acetyl-benzoyl)-2'-deoxycytidine, and  $N^4$ -(4-acetyl-benzoyl)-2'-deoxycytidine [16] was described in previous publications. The synthesis of other compounds is described in the Supplementary Material.

### 2.3. Growth Media

A standard Lennox LB medium was used for routine bacteria cultivation and protein synthesis induction (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl). SOB (20 g/L Bacto Tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 0.1 g/L KCl) was used for *E. coli* recovery after transformation. The complementation experiments with an auxotrophic *E. coli* HMS174

$\Delta pyrF \Delta yqfB$  strain were carried out in a minimal M9 medium (3.5 g/L  $\text{Na}_2\text{HPO}_4$ , 1.5 g/L  $\text{KH}_2\text{PO}_4$ , 2.5 g/L NaCl, 0.2 g/L  $\text{MgSO}_4$ , 0.01 g/L  $\text{CaCl}_2$ ) supplemented with 20 mg/L of ac4C.

For the cultivation of yeast cells, YPD (Yeast Peptone Dextrose; 1% yeast extract, 2% peptone, 2% dextrose) media with and without 2% agar was used for *P. pastoris* growth and storage. YPDS (YPD media with 1 M D-sorbitol) was used for *P. pastoris* cell recovery after transformation. PAD (Pichia Adenine Dropout media; 13.4% YNB-aa, 1.25% CMS-ADE, 0.005% biotin, 2% dextrose) with 2% agar was used for the growth and selection of *P. pastoris* transformants. BGMY (Buffered Glycerol-complex Medium; 1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% YNB-aa, 0.00004% biotin and 1% glycerol) and BMMY (Buffered methanol-complex medium; 1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% YNB-aa, 0.00004% biotin, and 0.5% methanol) was used for *P. pastoris* culture growth and induction.

#### 2.4. Gene Cloning into Expression Vectors

Genes encoding ASCH domain-containing proteins were PCR-amplified using the 2× Phusion Polymerase Master Mix (Thermo Scientific, Vilnius, Lithuania) and cloned into a pLATE31 expression vector by aLICator™ LIC Cloning and Expression System Kit 3 (Thermo Scientific, Vilnius, Lithuania). As a source of genomic DNA, intact cells of *B. agrestis*, *C. universalis*, *K. pneumoniae* KV-3, and *S. loihica* were used. Gene and primer sequences used in this study are listed in Tables S1 and S2. The cloning and expression procedures of the *E. coli yqfB* gene were described previously [7].

#### 2.5. Purification of YqfB Analogues

A single colony of *E. coli* HMS174  $\Delta pyrF \Delta yqfB$  cells transformed with the pLATE31 vector containing one of the target genes was inoculated into a 5 mL LB medium supplemented with ampicillin (50  $\mu\text{g}/\text{mL}$ ) and incubated at 37 °C with agitation (180 RPM) for 16 h. Overnight cultures were transferred into 50 mL of sterile LB medium and grown under the same conditions until  $\text{OD}_{600}$  reached 0.6–0.8. Protein synthesis was then induced by the addition of IPTG to a final concentration of 0.5 mM and incubated for an additional 12 h at 30 °C. The cells were harvested by centrifugation (10 min, 3220× *g*, 4 °C), resuspended in 20 mM Tris-HCl (pH 7.5) and 200 mM NaCl, and disrupted by sonication. The cell debris was removed by centrifugation (20 min, 16,000× *g*, 4 °C) and the supernatant was used for protein purification by the ÄKTA Pure protein purification system (Cytiva, Sweden). The cell-free extract was loaded on a 1 mL  $\text{Ni}^{2+}$ -charged chelating column (Cytiva, Sweden) and, using a buffer containing 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, and 500 mM imidazole, proteins were eluted by increasing the imidazole concentration. To remove the imidazole, protein samples were loaded on 2 × 5 mL desalting columns (Cytiva, Sweden) and eluted with 20 mM Tris-HCl (pH 7.5) and 200 mM NaCl buffer. Proteins were stored at −20 °C until usage. Protein purity was analyzed according to the standard Laemli SDS-PAGE protocol [17], then the concentration was measured using an Implen NanoPhotometer®.

#### 2.6. Determination of Substrate Range

Hydrolysis products of 12 modified cytosine and cytidine derivatives were determined qualitatively by applying thin-layer chromatography (TLC). The reaction mixture consisted of 4 mM of modified cytidine or 8 mM of modified cytosine compounds, 20 mM Tris-HCl (pH 7.5), and 1  $\mu\text{M}$  of protein of interest. After the incubation for 1 and 24 h in a 30 °C thermomixer, 1  $\mu\text{L}$  of the reaction mixture was spotted on a silica gel coated aluminum plate (60 F254, Merck, Germany), placed into a TLC chamber with chloroform:methanol (5:1 *v/v*) as an eluent. The dried plates were visualized with short wave UV light (254 nm).

The reaction mixture with the *p*-nitrophenyl acetate consisted of 1 mM of *p*-NP acetate, 20 mM Tris-HCl (pH 7.5), and 1  $\mu\text{M}$  of protein of interest. The absorption at 405 nm was measured spectrophotometrically after 30 min of incubation with an enzyme.

### 2.7. Enzyme Kinetics Assay

The activity of target amidohydrolases was measured by monitoring the hydrolysis of ac4C using the spectrophotometric method (Helios  $\gamma$  UV-Vis). The reaction mixture contained 20 mM Tris-HCl (pH 7.5), 0.025–0.2 mM of ac4C, and the appropriate amount of protein. Then, a decrease in absorbance at 310 nm was recorded for 30 s at 22 °C. The background hydrolysis of the substrate was determined by using a blank control with a composition identical to the reaction mixture, except the enzyme was replaced with a reaction buffer. All measurements were repeated at least three times. The kinetic parameters were determined by applying a non-linear regression model [18].

### 2.8. Relative Activity in Salts

The effect of salts on the enzyme activity was determined with NaCl and KCl concentrations of up to 4 M and 3 M, respectively. The reaction mixture contained 20 mM Tris-HCl (pH 7.5), 0.2 mM ac4C, and an appropriate amount of salt and protein. The hydrolytic activity was evaluated spectrophotometrically by measuring a decrease at  $A_{310}$ . The measurements were taken at 22 °C for 30 s.

### 2.9. Electrophoretic Mobility Shift Assay (EMSA)

Lyophilized *E. coli* MRE 600 tRNA (Roche, Germany) was dissolved in ultrapure MilliQ water treated with diethylpyrocarbonate (DEPC). The 5'-ends of the tRNA molecules were labeled with [ $\gamma$ - $^{32}$ P]-ATP using T4 polynucleotide kinase (Thermo Scientific, Lithuania). The 5'-labeled RNAs were incubated with target proteins for 30 min at 37 °C in a binding mixture, containing 20 mM Tris-HCl (pH 7.5), 300 mM KCl, 10 mM MgCl<sub>2</sub> or 10 mM CoCl<sub>2</sub>, and 100 ng/ $\mu$ L bovine serum albumin, similarly as described by Kim et al. [5]. The final ratio of tRNA to protein was 1:100. After the incubation, samples were loaded onto 8% (*v/v*) native polyacrylamide gel. Electrophoresis was conducted at 150 V for 150 min at 1 $\times$  TBE buffer. The visualization was performed by phosphorimaging (Fuji, Japan).

### 2.10. Cloning and Expression of Recombinant YqfB in *Pichia Pastoris*

The *yqfB* gene was amplified by PCR using the pET21b-*yqfB* plasmid as a template with the following primers: YqfB\_SmaI\_For (5'-TCCCCCGGGATGCAGCCAAACGACATCAC-3') and YqfB\_6xHis\_KpnI\_Rev (5'-GTCGGTAGGTACCTTAGTGGTGATGGTGATGATGAAGACATTTAAATTCATCAC-3'), where the highlighted parts are the restriction sites of SmaI and KpnI enzymes, respectively. The underlined part is the 6xHis-tag which is introduced at the C-terminus of YqfB. The resulting DNA fragment was cloned into the pPink- $\alpha$ F-HC (Invitrogen, Carlsbad, CA, USA) vector between the StuI and KpnI restriction sites. The constructed yeast expression vector pPink- $\alpha$ F-YqfB was verified by sequencing.

Transformation of *PichiaPink*<sup>TM</sup> strain 2 with pPink- $\alpha$ F-YqfB was performed according to the manufacturer's recommendations via electroporation. The plasmid DNA was digested with the SpeI restriction enzyme and later extracted with the phenol/chloroform mixture for further use. Five micrograms of linearized plasmid were used for the transformation. *PichiaPink*<sup>TM</sup> strains are *ade2* auxotrophs, so recombinant colonies were grown on PAD-Ag plates for an easy color-based selection of transformants. A single colony was inoculated in 12 mL of BGMY media in a 100 mL flask and grown at 30 °C in a shaking incubator for 2 days. The culture was used to inoculate 500 mL of BGMY media in 2 L baffled flasks and grown under the same conditions until OD<sub>600</sub> reached 5–6. The cells were then harvested by centrifugation at 3000 $\times$  *g* for 5 min. Protein expression was induced by resuspending the cells in 100 mL of BMMY media in 1 L baffled flasks. The culture was grown as described for another 72 h with the subsequent addition of methanol to a final concentration of 2% in 24-h intervals. The cells were pelleted by centrifugation at 3000 $\times$  *g* for 20 min, and the supernatant was collected. For efficient protein purification from the *P. pastoris* culture medium, the collected supernatant was placed into a dialysis bag and dialyzed against a phosphate buffer solution (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole; pH 8.0) at 4 °C for 48 h.

### 2.11. SDS-PAGE and Western Blot

Protein samples were mixed with a reducing sample buffer and heated to 95 °C prior to loading into a 14% SDS-polyacrylamide gel. Proteins were stained by the addition of Coomassie Brilliant Blue (Sigma-Aldrich Co.). After separation in SDS-PAGE, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Cytiva, Amersham PI, UK) by semi-dry electroblotting. The membrane was blocked with the Roti<sup>®</sup>Block protein-free blocking solution for 1 h (Carl Roth GmbH & Co. Kg, Karlsruhe, Germany), washed with PBS with 0.1% Tween-20 (PBS-T), and then incubated with primary 6xHis Tag Monoclonal antibodies (Mab) (HIS.H8) (Invitrogen, Carlsbad, CA, USA) 1:3000 diluted with PBS-T. After overnight incubation, the membrane was washed several times with PBS-T and subsequently incubated with goat anti-mouse IgG (H+L) HRP conjugate (Bio-Rad, Hercules, CA, USA) 1:3000 diluted with PBS-T. After washing, the chemiluminescence signal from the enzymatic reaction was developed by adding 4-chloro-1-naphthol and H<sub>2</sub>O<sub>2</sub> (Fluka, Buchs, Switzerland) or 1-Step<sup>™</sup> Ultra TMB-Blotting solution (Thermo Fisher Scientific, Vilnius, Lithuania).

### 2.12. Purification of YqfB from *P. pastoris* Growth Medium

Yeast-produced protein was purified by applying Ni<sup>2+</sup> affinity chromatography. Dialyzed yeast culture media were loaded onto a Ni<sup>2+</sup> immobilized column (Cytiva, Sweden), washed with a phosphate buffer solution, and eluted by increasing imidazole concentration. Then, gel filtration was applied to change the storage buffer to 20 mM Tris-HCl and 200 mM NaCl. The purified protein was concentrated using Amicon<sup>®</sup> Ultra-4 Ultracel<sup>®</sup>—10K centrifugal filter unit (Millipore, Ireland).

### 2.13. Glycosylation Analysis

Recombinant protein glycosylation was analyzed by an adapted Western blot (WB) assay. Instead of the primary antibodies, the membrane was incubated with *Canavalia ensiformis* (Jack bean) lectin Concanavalin A peroxidase conjugate (Sigma-Aldrich Co.) diluted with PBS-T to a working concentration of 0.1 µg/mL. After overnight incubation, the membrane was washed two times with PBS-T and the enzymatic reaction product was visualized by adding the 1-Step<sup>™</sup> Ultra TMB-Blotting solution (Thermo Scientific, Waltham, MA, USA).

## 3. Results

### 3.1. Amidohydrolytic Activity of YqfB Homologs

The potential proteins of interest for this study had to meet two main criteria: (1) sequence similarity to previously characterized YqfB from *E. coli* had to be from medium to high (between 50 and 90%), and (2) the availability of cultures in DMSZ and the possibility to grow them under available infrastructure. To do this, the phylogenetic tree was constructed using hits from the protein BLAST. Since *Klebsiella pneumoniae* isolate KV-3 (identity to YqfB—84%) was already present in the laboratory culture collection, the search was set to exclude *Escherichia* and *Klebsiella* genera from the initial query. Furthermore, the following parameters were applied: 50–80% sequence identity and query coverage of 95–100% to exclude truncated sequences. By performing this, in total, more than 1000 hits were obtained, meeting the abovementioned criteria. Based on the results of multiple sequence alignment, a phylogenetic tree was generated (Figure S1). For this study, four ASCH domain-containing proteins were chosen. These proteins were from mesophilic  $\gamma$ -proteobacteria *B. agrestis*, *C. universalis*, *K. pneumoniae*, and *S. loihica*, which fell into distinct phylogenetic groups based on the phylogenetic analysis results. Amplification of selected genes by PCR was followed by ligation independent cloning (LIC) into the pLATE31 expression vector. The amidohydrolytic activity of target proteins was first determined by transforming uridine auxotrophic *E. coli* HMS174  $\Delta$ pyrF $\Delta$ yqfB cells with plasmids encoding YqfB-like proteins and growing the cells on minimal media supplemented with 20 µg/mL ac4C and 50 µg/mL ampicillin (Figure S2). After overnight incubation, only

BagYqfB, CunYqfB, KpnYqfB, and SloYqfB-producing bacterial colonies appeared. Cells carrying pET21b-yqfB were used as a positive control, while cells transformed with the pET21b vector without an insert were used as a negative control. The ability of a uridine auxotrophic strain to grow on ac4C supplemented media was an indicator of the amidohydrolytic activity of the target protein [19].

Pairwise sequence alignment was used to compare the amino acid sequences of target enzymes with the previously described amidohydrolase YqfB. It was shown that KpnYqfB has the highest similarity (84%), followed by the less homologous BagYqfB (67%), CunYqfB (66%), and SloYqfB (52%) proteins. Additionally, the multiple sequence alignment (MSA) results showed that YqfB-like ASCH domain-containing proteins share identical amino acids at 39 positions. These include the catalytic residues Lys21, Thr24, Arg26, and Glu74 of YqfB (Figure S3a). Sequence visualization showed that, as expected, most of the amino acids that make up the structure of the ASCH domain were conserved among the proteins studied, while the variable regions were located in loops and on the surface of the enzymes (Figure S3b).

### 3.2. Substrate Specificity of Target Amidohydrolases

In order to qualitatively determine the substrate range of target proteins and to compare it with YqfB, purified enzymes were tested with  $N^4$ - and  $N^2$ -acylated cytosine as well as  $N^4$ -acylated 5'-fluorocytosine derivatives (Table 1). Samples were analyzed by performing TLC after 1 and 24 h of incubation and YqfB was used as a comparative control for this assay. Hydrolysis of a chromogenic substrate, *p*-nitrophenyl acetate, was monitored spectrophotometrically at 405 nm.

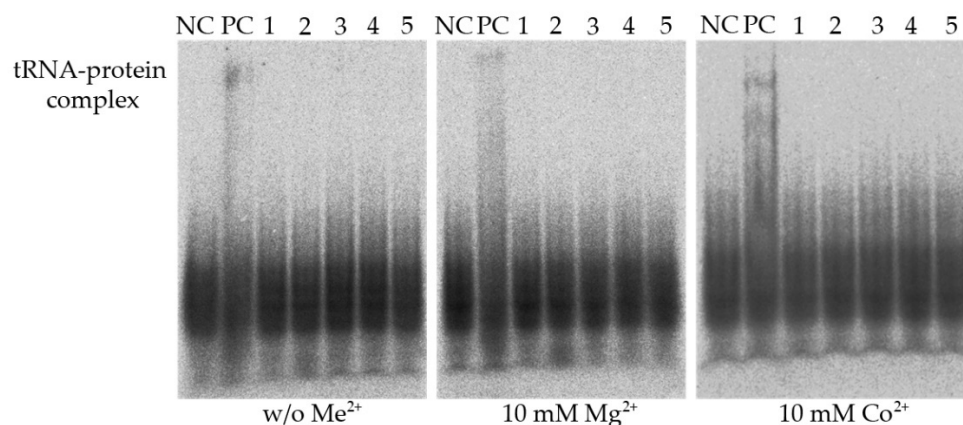
**Table 1.** A list of compounds used for qualitative determination of enzyme-substrate specificity. “+” indicates substrate hydrolysis within 24 h of incubation, “−”—hydrolysis did not occur.

Substrate	YqfB	BagYqfB	CunYqfB	KpnYqfB	SloYqfB
$N^4$ -acetylcytosine	+	+	+	+	+
$N^4$ -acetylcytidine	+	+	+	+	+
$N^4$ -benzoylcytidine	+	+	+	+	+
$N^4$ -acetyl-2'-deoxycytidine	+	+	+	+	+
$N^4$ -isobutyryl-2'-deoxycytidine	+	+	+	+	+
$N^4$ -hexanoyl-2'-deoxycytidine	+	+	+	+	+
$N^4$ -nicotinoyl-2'-deoxycytidine	+	+	+	+	+
$N^4$ -(3-acetyl-benzoyl)-2'-deoxycytidine	+	+	+	+	+
$N^4$ -(4-acetyl-benzoyl)-2'-deoxycytidine	+	+	+	+	+
$N^2$ -acetylisocytosine	+	+	+	+	+
$N^2$ -pivaloylisocytosine	−	−	−	−	−
Capecitabine	+	+	+	+	+
$N^2$ -acetyl-5-fluoroisocytosine	+	+	+	+	+
$N^4$ -benzoyl-5-fluorocytidine	+	+	+	+	+
$N^4$ -pivaloyl-5-fluorocytidine	+	+	+	+	+
$N^2$ -pivaloyl-5-fluoroisocytosine	−	−	−	−	−
<i>p</i> -nitrophenyl acetate	−	−	−	−	−

Despite the differences in amino acid sequences, their substrate range was identical to YqfB: all target proteins were active towards *N*-acylated (5'-fluoro)cytosine compounds,

harboring aliphatic, aromatic, and heterocyclic groups (Figure S5). Among 16 cytosine substrates tested, only two,  $N^2$ -pivaloylisocytosine and  $N^2$ -pivaloyl-5-fluoroisocytosine, were not hydrolyzed by target enzymes. This could be the result of the highly hydrophobic nature of the introduced modification. Furthermore, none of the enzymes were able to hydrolyse the *p*-NP ester. It should be noted that all enzymes recognized  $N^4$ -nicotinoyl-2'-deoxycytidine,  $N^4$ -(3-acetyl-benzoyl)-2'-deoxycytidine, and  $N^4$ -(4-acetyl-benzoyl)-2'-deoxycytidine; however, a significant amount of unhydrolyzed substrate was still detected even after 24 h incubation.

Additionally, regarding the homology of ASCH domain-containing proteins to pseudouridine synthase and archeosine transglycosylase domain (PUA) proteins [3], together with the recently published data on ZmASCH nucleic acid-binding and ssRNA ribonucleolytic activity [5], the *in vitro* electromobility shift assay (EMSA) was performed using Ni-NTA purified target amidohydrolases (Figure S4). *E. coli* tRNA and synthetic ssRNA (17 and 30 nucleotides length), as well as ssDNA (17 nucleotides), were tested as substrates for YqfB-like amidohydrolases, in both the absence and presence of divalent metal ions ( $Mg^{2+}$  or  $Co^{2+}$ ; data not shown). PUA domain-containing tRNA pseudouridine synthase B (TruB) was used for positive RNA binding control (Figure 1).



**Figure 1.** Evaluation of binding of YqfB and ASCH domain-containing YqfB-like proteins to tRNA in the presence and absence of metal ions in EMSA. NC—tRNA control without added protein samples, PC—positive control, tRNA pseudouridine synthase B (TruB). 1–5—tRNA incubated with YqfB, BagYqfB, CunYqfB, KpnYqfB and SloYqfB, respectively.

The EMSA results revealed no nucleic acid binding or nuclease activity of the target proteins under the tested conditions. However, in the case of the homologous domain harboring TruB, used as a positive control in this experiment, the formation of protein-tRNA complexes was not dependent on metal ions. This indicates that RNA molecules are unlikely to be the native substrates of the tested YqfB analogues.

### 3.3. Catalytic Properties of YqfB Analogues

Based on the TLC results, the target proteins were most active towards ac4C as well as ac4dC (substrates were completely hydrolyzed within 1 h). Referring to the previously published YqfB data, we suggest that ac4C could be the primary substrate for the target proteins in the cell. Following this, the kinetics of the recombinant enzyme were analyzed (Table 2).

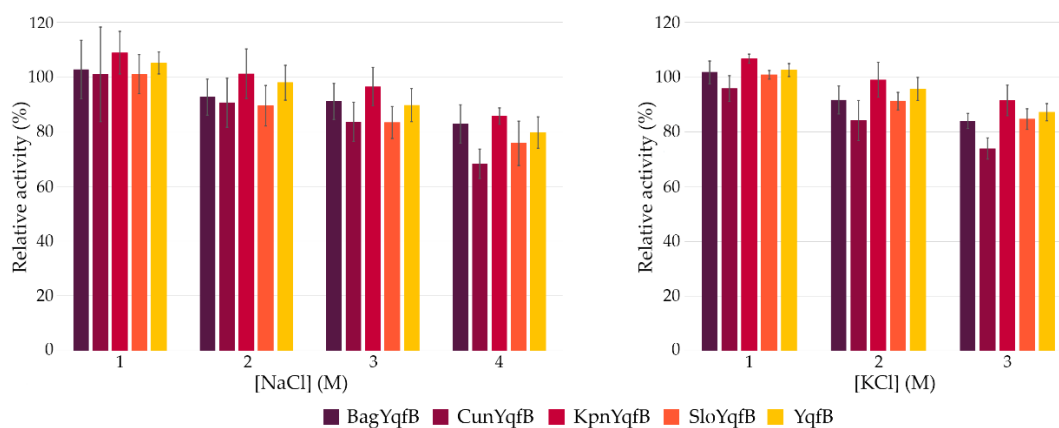
Similar kinetic parameters indicate that ac4C could be the primary substrate for all four investigated ASCH domain-containing proteins, which possessed a two-fold higher catalytic efficiency compared to YqfB. The largest difference appeared in the case of SloYqfB, which showed a lower affinity for the substrate, which was compensated for by the highest reaction rate among the tested enzymes. It should also be noted that a possible inhibition was not taken into account in the calculations. We assume that substrate hydrolysis occurs under the same acid–base mechanism as described for YqfB [7]. The unique catalytic triad,

found at the domain's conservative octapeptide, consists of Thr and Lys, which act as a nucleophile and a base, respectively, while Glu is an acid.

**Table 2.** Kinetic parameters of YqfB and ASCH domain-containing YqfB-like proteins. The activity was measured in 20 mM Tris-HCl, pH 7.5, at 22 °C, using ac4C as a substrate.

Enzyme	$K_M$ (M)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $M^{-1}s^{-1}$ )
YqfB [7]	$(6.2 \pm 0.1) \times 10^{-5}$	$157 \pm 1$	$(2.2 \pm 0.1) \times 10^6$
BagYqfB	$(3.6 \pm 0.8) \times 10^{-5}$	$154 \pm 11$	$(4.3 \pm 0.7) \times 10^6$
CunYqfB	$(5.7 \pm 0.8) \times 10^{-5}$	$248 \pm 20$	$(4.4 \pm 0.5) \times 10^6$
KpnYqfB	$(4.6 \pm 0.8) \times 10^{-5}$	$216 \pm 2$	$(4.8 \pm 0.8) \times 10^6$
SloYqfB	$(1 \pm 0.2) \times 10^{-4}$	$441 \pm 37$	$(4.3 \pm 0.5) \times 10^6$

Unexpectedly, a high level of activity of the YqfB-like enzymes in high molarity of salts (NaCl, KCl) was determined (Figure 2).



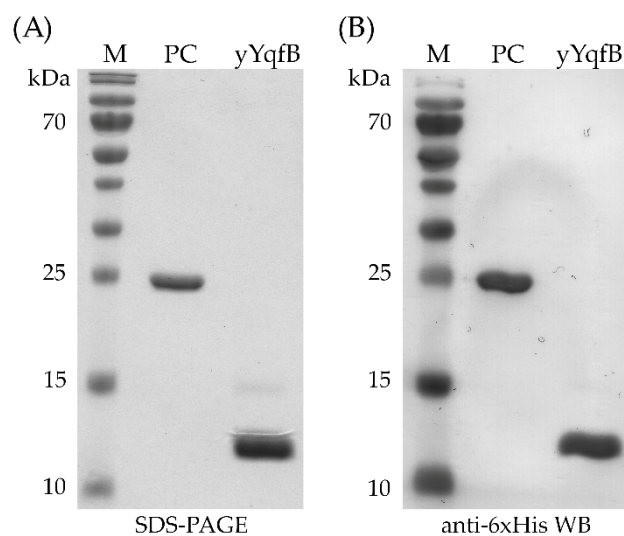
**Figure 2.** Effect of NaCl and KCl on the activity of ASCH domain-containing proteins. The activity without added salts was taken as 100%. The substrate used in these experiments was ac4C.

All tested YqfB analogues showed a very high salt tolerance and only a slight (~20%) decrease in activity at extreme concentrations of KCl and NaCl was observed. From all tested proteins, KpnYqfB and CunYqfB possessed the highest and lowest tolerance to high salt concentrations, respectively, although the differences did not reach statistical significance.

### 3.4. YqfB Biosynthesis in a Yeast Expression System

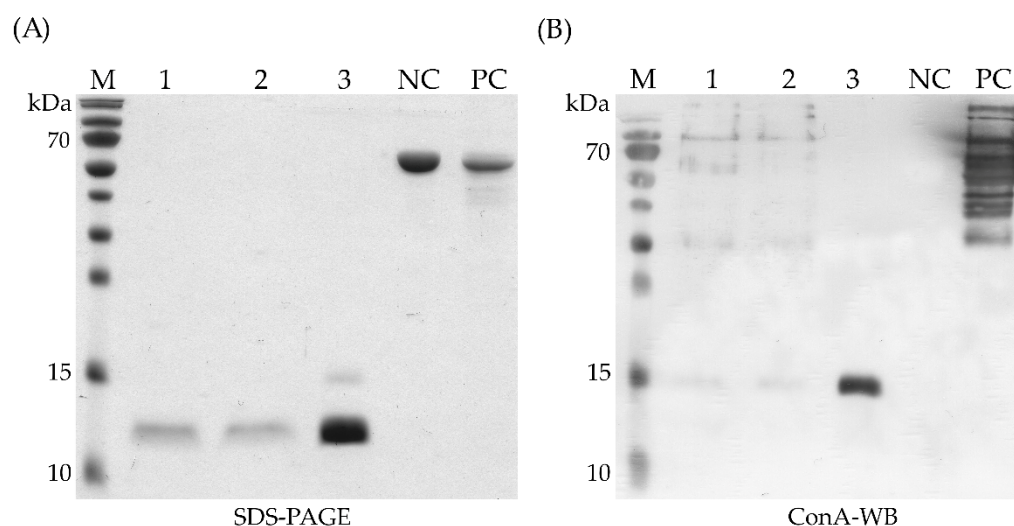
Considering the possibility of using small amidohydrolases containing ASCH domains for therapeutic purposes, a GRAS-approved eukaryotic *P. pastoris* expression system was chosen as an alternative host. By adding a secretion sequence to YqfB, we were able to purify the enzyme from the culture medium. The SDS-PAGE results revealed two bands in the cell growth medium sample—the target YqfB band at ~12 kDa and another at ~15 kDa. After Ni-NTA purification, protein profile analysis revealed that the ~15 kDa protein elutes with YqfB (Figure 3a). The anti-6xHis Western blot analysis showed a target protein band at 12 kDa along with a low-intensity band at 15 kDa. This suggests that in yeast cells, a secretion sequence harboring YqfB could undergo post-translational modifications (Figure 3b).





**Figure 3.** Analysis of YqfB produced in yeast *P. pastoris* (yYqfB). (A) Protein profile in SDS-PAGE, where M is the PageRuler Prestained Protein Ladder; PC is the 6xHis-tagged *Penaeus monodon* allergen Pen m4 –positive control for anti-6xHis Western blot analysis; yYqfB is the yeast produced *E. coli* YqfB; (B) Western blot analysis of described samples.

To investigate, whether part of the bacterial protein is glycosylated in the yeast expression system, a Western blot with concanavalin A (ConA) was performed. ConA is known as a lectin that specifically binds  $\alpha$ -D-mannosyl and  $\alpha$ -D-glucosyl groups, and is used to detect glycosylated proteins [20]. The results showed a similar band at ~15 kDa, while the main portion of the protein was ConA negative (Figure 4b). YqfB amino acid sequence analysis revealed a single position where *N*-glycosylation could occur—Asn76 (see Figure S6) [21].



**Figure 4.** Determination of *N*-glycosylation of YqfB after expression in *P. pastoris* cells. (A) Protein profiles in SDS-PAGE and (B) Western Blot analysis with ConA. 1–3—protein samples prior to (1) and after (2) dialysis of yeast growth media and purification by affinity chromatography (3). PC—*P. pastoris*-produced glycosylated black tiger shrimp protein, Pen m 4, fused to the Maltose-binding protein (MBP), acts as a positive control for Western blot analysis with ConA. NC—*E. coli*-produced Pen m 4 protein fused to MBP acts as a negative control.

Qualitative enzyme activity experiments with ac4C showed that the yeast-produced protein retained its amidohydrolytic activity. The ability to produce a secreted active form of the target enzyme is important for a more convenient downstream process in biotechnology.

#### 4. Discussion

In this work, we characterized four new ASCH domain-containing amidohydrolases from the mesophilic bacteria *B. agrestis*, *C. universalis*, *K. pneumoniae* KV-3, and *S. loihica*. The target proteins share moderate to high (52–84 %) sequence similarity to *E. coli* YqfB. Multiple sequence alignment revealed that catalytic Lys21, Thr24, and Glu74 amino acids, involved in the hydrolysis of modified cytidine compounds, are shared by all of the enzymes described in this study. According to the predicted 3D models, most of the conserved amino acids make up the core structure of the ASCH domain, while the variable regions are located in loops and on the enzyme surface. The above-mentioned catalytic triad found in BagASCH, CunASCH, KpnASCH, and SloASCH along with YqfB is an atypical and unique feature of this group of amidohydrolases [7]. The enzymes exhibited broad substrate specificity with *N*-acylated cytosine compounds. Based on the available enzymatic reaction parameters, we assume that in the cell, the main substrate of these enzymes is ac4C, since the calculated catalytic efficiency ( $k_{\text{cat}}/K_M$ ) of all target enzymes is  $\sim 4.5 \times 10^6$ . In addition to their small size and high catalytic activity, the target proteins retain  $\sim 80$  % of their activity in 3 M KCl and 4 M NaCl. The significant halotolerance of these small amidohydrolases could be the result of an acidic amino acid-rich protein sequence [22,23]. For the studied YqfB analogues, Glu and Asp residues constitute up to 20 % of the amino acids in the sequence.

It is known that the number of ac4C positions in RNA molecules increases when the cell is exposed to environmental stress, which would suggest that the introduction of this modification is another example of dynamic adaptation [24,25]. However, there is no evidence for the reversibility of this modification in the bacteria domain. The high specificity of YqfB and its analogues to ac4C as well as the published data on ZmASCH ribonuclease activity have led to the hypothesis that these small amidohydrolases could carry out RNA deacetylation in vivo or be involved in the degradation of acetylated RNAs [5]. However, since no nucleic acid binding or ribonuclease activity was detected, the high catalytic activity of the tested enzymes suggests that free acetylated cytidine nucleotides are the cellular substrates rather than RNAs. This points to a possible role of ASCH domain-containing proteins in cytidine metabolism, since removing the acyl-group from modified cytidine could return the nucleotide to the pyrimidine salvage pathway. It is also possible that RNA-YqfB binding in vivo requires additional docking proteins, which were not analyzed in this study.

The ability of YqfB-like proteins to remove cytidine modifications is not only of fundamental importance but also of great interest for practical applications. In this study, we showed that all target proteins are capable of hydrolyzing capecitabine and other 5-fluorouracil-derived compounds. This suggests the possible role of ASCH domain-containing proteins in the prodrug activating systems, where an inactive prodrug is enzymatically converted into a cancer cells-targeting molecule. Considering the halotolerance of the YqfB-like enzymes, increased stability of these proteins in the blood could also be suggested. Although the bacterial expression system is efficient for the biosynthesis of these small amidohydrolases, a GRAS organism should be considered for potential application in pharmaceuticals. Here we show that the addition of a secretion signal sequence to the *N*-terminus of *E. coli* YqfB led to successful functional enzyme synthesis and transport to the extracellular environment by employing a well-studied eukaryotic expression system, *P. pastoris*, thus expanding the potential applications of these small amidohydrolases.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12101492/s1>, description of: Synthesis of N4-modified nucleobases and nucleosides; Table S1: DNA and amino acid sequences of YqfB-type proteins used in this study; Table S2: List of primers used in this study; Figure S1: Phylogenetic tree of ASCH domain-containing proteins; Figure S2: Growth of *E. coli* HMS174  $\Delta$ pyrF  $\Delta$ yqfB cells on minimal media; Figure S3: Analysis of conserved regions of YqfB-type amidohydrolases; Figure S4: Profiles of purified YqfB analogues in SDS-PAGE; Figure S5: Qualitative enzymatic activity analysis of amidohydrolases; Figure S6: Predicted YqfB N-glycosylation sites in yeast using NetNGlyc—1.0 server.

**Author Contributions:** Conceptualization, R.S. and R.M.; methodology, R.S., M.S., K.K. and J.R.; investigation, R.S., M.S., K.K. and J.R.; resources, R.M.; writing—original draft preparation, R.S. and J.R.; writing—review and editing, R.S., M.S., J.R. and R.M.; visualization, R.S. and J.R.; supervision, R.M.; funding acquisition, R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project has received funding from European Regional Development Fund (project No. 01.2.2-LMT-K-718-03-0082) under grant agreement with the Research Council of Lithuania (LMTLT).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sprinzl, M.; Horn, C.; Brown, M.; Ioudovitch, A.; Steinberg, S. Compilation of TRNA Sequences and Sequences of TRNA Genes. *Nucleic Acids Res.* **1998**, *26*, 148–153. [[CrossRef](#)] [[PubMed](#)]
2. Jung, D.-J.; Sung, H.-S.; Goo, Y.-W.; Lee, H.M.; Park, O.K.; Jung, S.-Y.; Lim, J.; Kim, H.-J.; Lee, S.-K.; Kim, T.S.; et al. Novel Transcription Coactivator Complex Containing Activating Signal Cointegrator 1. *Mol. Cell. Biol.* **2002**, *22*, 5203–5211. [[CrossRef](#)] [[PubMed](#)]
3. Iyer, L.M.; Burroughs, A.M.; Aravind, L. The ASCH Superfamily: Novel Domains with a Fold Related to the PUA Domain and a Potential Role in RNA Metabolism. *Bioinformatics* **2006**, *22*, 257–263. [[CrossRef](#)] [[PubMed](#)]
4. Bertonati, C.; Punta, M.; Fischer, M.; Yachdav, G.; Forouhar, F.; Zhou, W.; Kuzin, A.P.; Seetharaman, J.; Abashidze, M.; Ramelot, T.A.; et al. Structural Genomics Reveals EVE as a New ASCH/PUA-Related Domain. *Proteins* **2009**, *75*, 760–773. [[CrossRef](#)] [[PubMed](#)]
5. Kim, B.-N.; Shin, M.; Ha, S.C.; Park, S.-Y.; Seo, P.-W.; Hofmann, A.; Kim, J.-S. Crystal Structure of an ASCH Protein from *Zymomonas Mobilis* and Its Ribonuclease Activity Specific for Single-Stranded RNA. *Sci. Rep.* **2017**, *7*, 12303. [[CrossRef](#)] [[PubMed](#)]
6. Park, S.-Y.; Park, J.-H.; Kim, J.-S. Cloning, Expression, Purification, Crystallization and Preliminary X-ray Diffraction Analysis of an ASCH Domain-Containing Protein from *Zymomonas Mobilis* ZM4. *Acta Cryst. Sect. F Struct. Biol. Cryst. Commun.* **2011**, *67*, 310–312. [[CrossRef](#)] [[PubMed](#)]
7. Stanislauskienė, R.; Laurynėnas, A.; Rutkienė, R.; Aučynaitė, A.; Tauraitė, D.; Meškienė, R.; Urbelienė, N.; Kaupinis, A.; Valius, M.; Kaliniene, L.; et al. YqfB Protein from *Escherichia coli*: An Atypical Amidohydrolase Active towards N4-Acylcytosine Derivatives. *Sci. Rep.* **2020**, *10*, 788. [[CrossRef](#)] [[PubMed](#)]
8. Arango, D.; Sturgill, D.; Alhusaini, N.; Dillman, A.A.; Sweet, T.J.; Hanson, G.; Hosogane, M.; Sinclair, W.R.; Nanan, K.K.; Mandler, M.D.; et al. Acetylation of Cytidine in mRNA Promotes Translation Efficiency. *Cell* **2018**, *175*, 1872–1886.e24. [[CrossRef](#)] [[PubMed](#)]
9. Ito, S.; Akamatsu, Y.; Noma, A.; Kimura, S.; Miyauchi, K.; Ikeuchi, Y.; Suzuki, T.; Suzuki, T. A Single Acetylation of 18 S RRNA Is Essential for Biogenesis of the Small Ribosomal Subunit in *Saccharomyces Cerevisiae*. *J. Biol. Chem.* **2014**, *289*, 26201–26212. [[CrossRef](#)] [[PubMed](#)]
10. Kumbhar, B.V.; Kamble, A.D.; Sonawane, K.D. Conformational Preferences of Modified Nucleoside N(4)-Acetylcytidine, Ac4C Occur at “Wobble” 34th Position in the Anticodon Loop of TRNA. *Cell. Biochem. Biophys.* **2013**, *66*, 797–816. [[CrossRef](#)]
11. Wada, T.; Kobori, A.; Kawahara, S.; Sekine, M. Synthesis and Properties of Oligodeoxyribonucleotides Containing 4-N-Acetylcytosine Bases. *Tetrahedron Lett.* **1998**, *39*, 6907–6910. [[CrossRef](#)]
12. Kawai, G.; Hashizume, T.; Miyazawa, T.; McCloskey, J.A.; Yokoyama, S. Conformational Characteristics of 4-Acetylcytidine Found in TRNA. *Nucleic Acids Symp Ser.* **1989**, *21*, 61–62.
13. Madeira, F.; Pearce, M.; Tivey, A.R.N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50*, gkac240. [[CrossRef](#)] [[PubMed](#)]
14. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)] [[PubMed](#)]
15. Šimoliūnas, E.; Kaliniene, L.; Truncaitė, L.; Zajančauskaitė, A.; Staniulis, J.; Kaupinis, A.; Ger, M.; Valius, M.; Meškys, R. Klebsiella Phage VB\_KleM-RaK2—A Giant Singleton Virus of the Family Myoviridae. *PLoS ONE* **2013**, *8*, e60717. [[CrossRef](#)]
16. Jakubovska, J.; Tauraitė, D.; Birštonas, L.; Meškys, R. N4-Acyl-2'-Deoxycytidine-5'-Triphosphates for the Enzymatic Synthesis of Modified DNA. *Nucleic Acids Res.* **2018**, *46*, 5911–5923. [[CrossRef](#)]
17. Laemmli, U.K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227*, 680–685. [[CrossRef](#)]
18. Nonlinear Regression Michaelis & Menten. Available online: <http://biomodel.uah.es/en/metab/enzimas/MM-regresion.htm> (accessed on 28 June 2022).

19. Urbelienė, N.; Meškienė, R.; Tiškus, M.; Stanislauskienė, R.; Aučynaitė, A.; Laurynėnas, A.; Meškys, R. A Rapid Method for the Selection of Amidohydrolases from Metagenomic Libraries by Applying Synthetic Nucleosides and a Uridine Auxotrophic Host. *Catalysts* **2020**, *10*, 445. [[CrossRef](#)]
20. COULIBALY, F.S.; YOUAN, B.-B.C. Concanavalin A—Polysaccharides Binding Affinity Analysis Using A Quartz Crystal Microbalance. *Biosens. Bioelectron.* **2014**, *59*, 404–411. [[CrossRef](#)]
21. Gupta, R.; Brunak, S. Prediction of Glycosylation across the Human Proteome and the Correlation to Protein Function. In *Pacific Symposium on Biocomputing 2002, Proceedings of the Pacific Symposium, Kauai, Hawaii, USA, 3–7 January 2002*; Altman, R.B., Dunker, A.K., Hunter, L., Klein, T.E., Eds.; World Scientific: Singapore, 2002; pp. 310–322.
22. Ebrahimie, E.; Ebrahimi, M.; Sarvestani, N.R.; Ebrahimi, M. Protein Attributes Contribute to Halo-Stability, Bioinformatics Approach. *Saline Syst.* **2011**, *7*, 1. [[CrossRef](#)]
23. Paul, S.; Bag, S.K.; Das, S.; Harvill, E.T.; Dutta, C. Molecular Signature of Hypersaline Adaptation: Insights from Genome and Proteome Composition of Halophilic Prokaryotes. *Genome Biol.* **2008**, *9*, R70. [[CrossRef](#)] [[PubMed](#)]
24. Tardu, M.; Jones, J.D.; Kennedy, R.T.; Lin, Q.; Koutmou, K.S. Identification and Quantification of Modified Nucleosides in *Saccharomyces Cerevisiae* MRNAs. *ACS Chem. Biol.* **2019**, *14*, 1403–1409. [[CrossRef](#)] [[PubMed](#)]
25. Orita, I.; Futatsuishi, R.; Adachi, K.; Ohira, T.; Kaneko, A.; Minowa, K.; Suzuki, M.; Tamura, T.; Nakamura, S.; Imanaka, T.; et al. Random Mutagenesis of a Hyperthermophilic Archaeon Identified TRNA Modifications Associated with Cellular Hyperthermotolerance. *Nucleic Acids Res.* **2019**, *47*, 1964–1976. [[CrossRef](#)] [[PubMed](#)]