



Article

Rational Mutagenesis in the Lid Domain of Ribokinase from *E. coli* Results in an Order of Magnitude Increase in Activity towards D-arabinose

Evgeniy A. Zayats * , Ilya V. Fateev , Maria A. Kostromina, Yulia A. Abramchik , Dmitry D. Lykoshin , Daria O. Yurovskaya, Vladimir I. Timofeev, Maria Ya. Berzina , Barbara Z. Eletskaia, Irina D. Konstantinova * and Roman S. Esipov

Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya St. 16/10, 117997 Moscow, Russia

* Correspondence: eaz96post@gmail.com (E.A.Z.); kid1968@yandex.ru (I.D.K.)

Abstract: Development of efficient approaches for the production of medically important nucleosides is a highly relevant challenge for biotechnology. In particular, cascade synthesis of arabinosides would allow relatively easy production of various cytostatic and antiviral drugs. However, the biocatalyst necessary for this approach, ribokinase from *Escherichia coli* (*EcoRK*), has a very low activity towards D-arabinose, making the synthesis using the state-of-art native enzyme technologically unfeasible. Here, we report the results of our enzyme design project, dedicated to engineering a mutant form of *EcoRK* with elevated activity towards arabinose. Analysis of the active site structure has allowed us to hypothesize the reasons behind the low *EcoRK* activity towards arabinose and select feasible mutations. Enzyme assay and kinetic studies have shown that the A98G mutation has caused a large 15-fold increase in k_{cat} and 1.5-fold decrease in K_M for arabinose phosphorylation. As a proof of concept, we have performed the cascade synthesis of 2-chloroadenine arabinoside utilizing the A98G mutant with 10-fold lower amount of enzyme compared to the wild type without any loss of synthesis efficiency. Our results are valuable both for the development of new technologies of synthesis of modified nucleosides and providing insight into the structural reasons behind *EcoRK* substrate specificity.

Keywords: enzyme genetic improvement; rational design; active site modification; site-directed mutagenesis; ribokinase; cascade synthesis; modified nucleoside; arabinoside; arabinose



Citation: Zayats, E.A.; Fateev, I.V.; Kostromina, M.A.; Abramchik, Y.A.; Lykoshin, D.D.; Yurovskaya, D.O.; Timofeev, V.I.; Berzina, M.Y.; Eletskaia, B.Z.; Konstantinova, I.D.; et al. Rational Mutagenesis in the Lid Domain of Ribokinase from *E. coli* Results in an Order of Magnitude Increase in Activity towards D-arabinose. *Int. J. Mol. Sci.* **2022**, *23*, 12540. <https://doi.org/10.3390/ijms232012540>

Academic Editor: Roberto Fernandez-Lafuente

Received: 7 September 2022

Accepted: 16 October 2022

Published: 19 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/>).

1. Introduction

One of the key challenges of modern biotechnology is the development of efficient approaches for the synthesis of medically important drugs. A noteworthy class of compounds that are broadly applied for treatment of a large variety of diseases are modified nucleosides [1]. These compounds act as mimics of their natural nucleoside and nucleotide counterparts, therefore interfering in the biological processes that are dependent on DNA and RNA metabolism [2]. Such biological activity allows these compounds to suppress the proliferation of tumor cells, bacteria and viruses, therefore resulting in broad-spectrum activity against cancer and infectious diseases. Therefore, the development of new and efficient technologies for the production of modified nucleosides is an important scientific task.

Conventional approaches involving chemical synthesis remain a laborious multi-stage process requiring the introduction and removal of protective groups, in some cases resulting in poor product yield and limited process efficiency [1,3]. As an alternative, cascades of enzymes involved in nucleic acid metabolism offer single-step one-pot synthesis of nucleosides and nucleotides with high stereo- and regioselectivity [1,3–6]. Nevertheless,

one of the key factors limiting the efficiency of the cascade synthesis of modified nucleosides is the low activity of the wild type enzymes towards the necessary unnatural substrates.

On one hand, enzymes that carry out the glycosylation reaction in the cascade synthesis of nucleosides tend to have broad substrate specificity towards the nucleobase substrates [7–9]. On the other hand, enzymatic synthesis of nucleosides with a modified sugar moiety turns out to be more challenging. Sugar-modified nucleosides availability is strictly limited. Therefore, these compounds cannot be obtained from commodity reagents through nucleoside phosphorylase mediated transglycosilation. To overcome this challenge, a three-enzyme (ribokinase, phosphopentomutase and nucleoside phosphorylase) cascade was developed, offering a synthesis starting from easily accessible pentoses (Figure 1) [4,10].

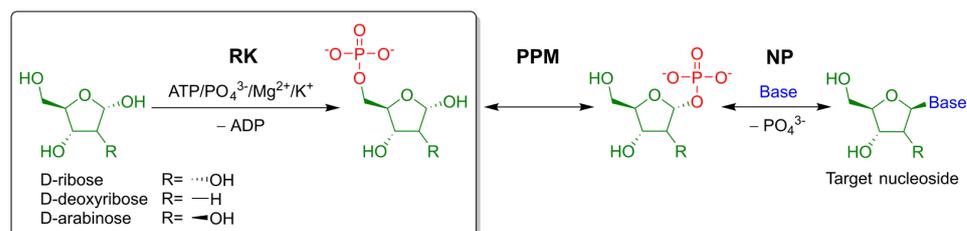


Figure 1. A three-enzyme cascade for the synthesis of modified nucleosides. RK—ribokinase (ec:2.7.1.15), PPM—phosphopentomutase (ec: 5.4.2.7), NP—nucleoside phosphorylase, usually purine nucleoside phosphorylase (ec:2.4.2.1). Base—nucleobase substrate/nucleoside moiety.

The first enzyme of this cascade is ribokinase (ec:2.7.1.15). This enzyme catalyzes the phosphorylation of D-ribose through the 5'-hydroxyl group, using ATP as the phosphate donor (Figure 1, first reaction). Ribokinase from both *E. coli* and other sources is a homodimer with two symmetrical active sites [11–14]. Each subunit contains two functional domains: the core $\alpha + \beta$ fold and the so-called lid domain (Figure 2A) [11]. The lid domain is a five-stranded β -sheet, formed through cooperation between enzyme subunits: four strands come from one subunit and one from another.

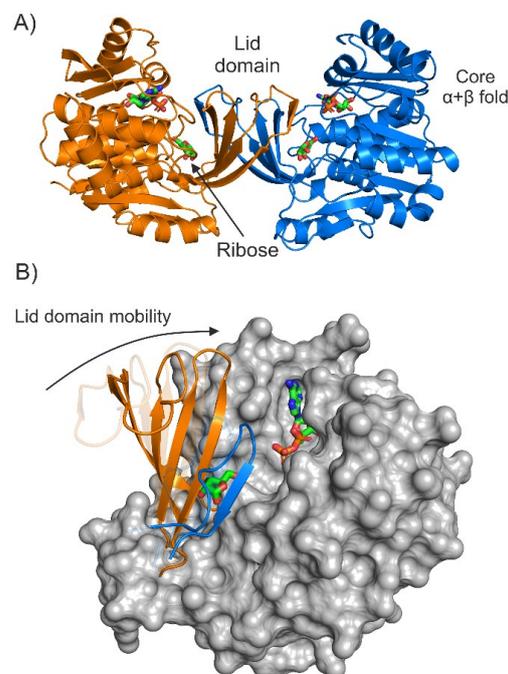


Figure 2. (A)—The crystal structure of the *EcoRK* dimer (PDB 1RKD). (B)—The mobility of the lid domain during the transition from APO form (semi-transparent, spatial alignment with crystal structure of APO *EcoRK*, PDB 1RKA) to substrate-bound form (PDB 1RKD). The lid domain is represented as cartoon, all other residues—as a surface.

The transition of *EcoRK* from APO to a substrate-bound form is accompanied by a significant conformational change due to the lid-domain mobility. The mentioned 5-stranded β -sheet acts as a lid (hence the name) that is open in APO form and closed upon ribose binding (Figure 2B) [11,12]. Ribose binding occurs through a network of hydrogen bonds between the hydroxyl carbohydrate groups and the active site residues: Asn14, Asp16, Gly42, Lys43, Asn46, Glu143 and Asp255. A residue interacting with the 5'-hydroxyl group of ribose, Asp255 is believed to be catalytic and responsible for the deprotonation of this group, preceding the formation of an ester with a phosphate [11].

The ability of ribokinase to phosphorylate pentoses allows the introduction of carbohydrates (including those that are not present as moieties of natural nucleosides) into enzymatic cascade-reaction pathways. Therefore, utilization of this enzyme is a valuable alternative to transglycosylation for the synthesis of nucleosides with unnatural pentose moieties. Unfortunately, the substrate specificity of ribokinase is relatively narrow. We have previously shown that for both ribokinase from *Escherichia coli* (*EcoRK*) and *Thermus sp. 2.9* (*TspRK*), only ribose and deoxyribose are good substrates [15,16].

One of the most medically important classes of sugar-modified nucleoside analogues are the derivatives of D-arabinose. Noteworthy broadly applied arabinoside drugs include compounds with both cancer suppressing (fludarabine, nelarabine, cytarabine) and antiviral (vidarabine) activity [17–20]. Our previous studies show that arabinose is a substrate for both *EcoRK* and *TspRK*, although the enzymatic activity towards this pentose is very low [4,15,16]. According to Park et al., the human ribokinase has a very weak activity towards arabinose as well [21]. We have previously shown that the wild type ribokinases from both *E. coli* and *T. sp.* can be used for cascade synthesis of arabinonucleosides [4,22]. However, even at a very high excess of arabinose, a large amount of the recombinant enzyme was still necessary for efficient synthesis. Therefore, the feasibility of industrial cascade synthesis of arabinonucleoside drugs using the state-of-art wild type enzymes remains highly questionable.

To summarize, the narrow carbohydrate substrate specificity of ribokinase limits its applicability for the synthesis of a variety of medically important arabinonucleoside drugs. A rapidly developing approach, offering an opportunity to overcome this challenge is enzyme design. An absent but desired property of a biocatalyst (in our case, substrate promiscuity) can be engineered through the study of the enzyme active site and subsequent rational modification via site-directed mutagenesis. In our research piece, we report the successful rational design of a mutant form of ribokinase from *Escherichia coli* with a significant increase in turnover rate (k_{cat}) and decrease in K_M . We have chosen this enzyme instead of *TspRK* since *EcoRK* has a broad temperature range of activity, allowing to design cascades with enzymes from both mesophilic and thermophilic sources [4]. Our research is significant both for developing new technologies for the cascade synthesis of medically important compounds and for the study of the relationship between ribokinase active site structure and function.

2. Results

2.1. Selection of Design Hot Spots and Feasible Mutations

The first stage of our enzyme design project was to study the structure of the active site of ribokinase from *Escherichia coli* in order to hypothesize the factors determining its substrate specificity, select hotspots for enzyme design and choose mutations beneficial for arabinose binding.

Our compound of interest, arabinose, is a stereoisomer of the natural substrate ribose. The only difference between these molecules is the chirality of the 2' carbon ($C2'$) atom. Therefore, if we inverse the chirality of this atom of ribose in the crystal structure of *EcoRK*, we can make an initial estimation of the position of arabinose in the active site. On Figure 3, we can observe that ribose assumes the $O4'$ -endo conformation in the active site of *EcoRK* (PDB 1RKD). We could hypothesize that arabinose binds to *EcoRK* in a way that is similar

to the natural substrate. In order to verify this assumption, we have studied the binding of ribose and arabinose by *Eco*RK in silico.

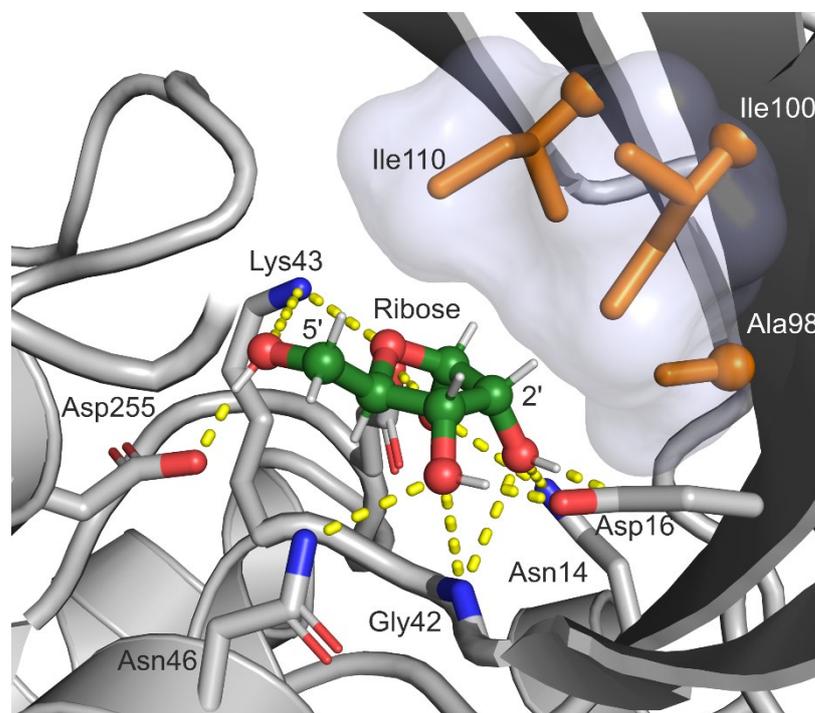


Figure 3. The position of ribose in the active site of *Eco*RK (PDB 1RKD). The semi-transparent surface represents the hydrophobic interface in the lid domain, formed by Ala98, Ile100 and Ile110. Residues 248–251 and 285–295 are hidden. Carbon atoms of ribose are colored green, carbon atoms of residues Ala98, Ile100 and Ile110—orange, oxygen atoms—red, nitrogen atoms—blue. Yellow dashed lines represent polar contacts. Visualization was performed in PyMOL 2.0 (Schrödinger, LLC).

To perform a comparative study of ribose and arabinose binding in the active site of *Eco*RK, we have carried out molecular dynamics (MD) simulations in Gromacs 2021 using the Amber AMBER99SB-ILDN forcefield. The structure of *Eco*RK in complex with the natural substrate was fetched from Protein Data Bank (PDB 1RKD). Complex with arabinose (initial approximation) was obtained by reversing the chirality of the C2' ribose atom through molecular editing. Rhombic dodecahedral periodic cells with protein-ligand complexes were filled with tip3p water-molecule models, neutralized with 150 mM KCl ions and subjected to energy minimization, followed by NVT and NPT equilibration. The final step was productive MD, which resulted in two 400-ns long MD simulation trajectories.

An analysis of the obtained trajectories (RMSD, RMSF and radius of gyration) verified the stability of the complex between *Eco*RK and both the natural and unnatural substrate throughout the MD simulations (Figure 4A–D). Therefore, the mentioned MD simulation trajectories are appropriate for further analysis.

The conformation of the active site of *Eco*RK in complex with arabinose (obtained from cluster analysis) is illustrated in Figure 5. We can observe that most of the interactions between arabinose and the active site are similar to those of the natural substrate (Figure 3). However, calculation of furanose ring pseudorotation coordinates via PLUMED has revealed that arabinose assumes the $\frac{3}{2}$ Twist puckering conformation, unlike the O4'-endo envelope of ribose (Figure 4E). Therefore, the experiment in silico has shown that while the interactions of ribose and arabinose with active site residues are similar, the conformation of the ligand itself is different.

One of the main differences between ribose and arabinose interactions with the active site residues is that the Asn14 residue does not form a hydrogen bond with arabinose, which may be a factor hindering *Eco*RK activity towards the unnatural substrate. Nevertheless,

according to Kang et al., mutation of a homologous Asn78 residue to alanine in ribokinase from *Arabidopsis thaliana* did not result in a significant decrease in activity towards ribose [23]. This fact suggests that the loss of aforementioned hydrogen bond may not be critical for pentose phosphorylation. Therefore, investigation of other factors is a more feasible research direction.

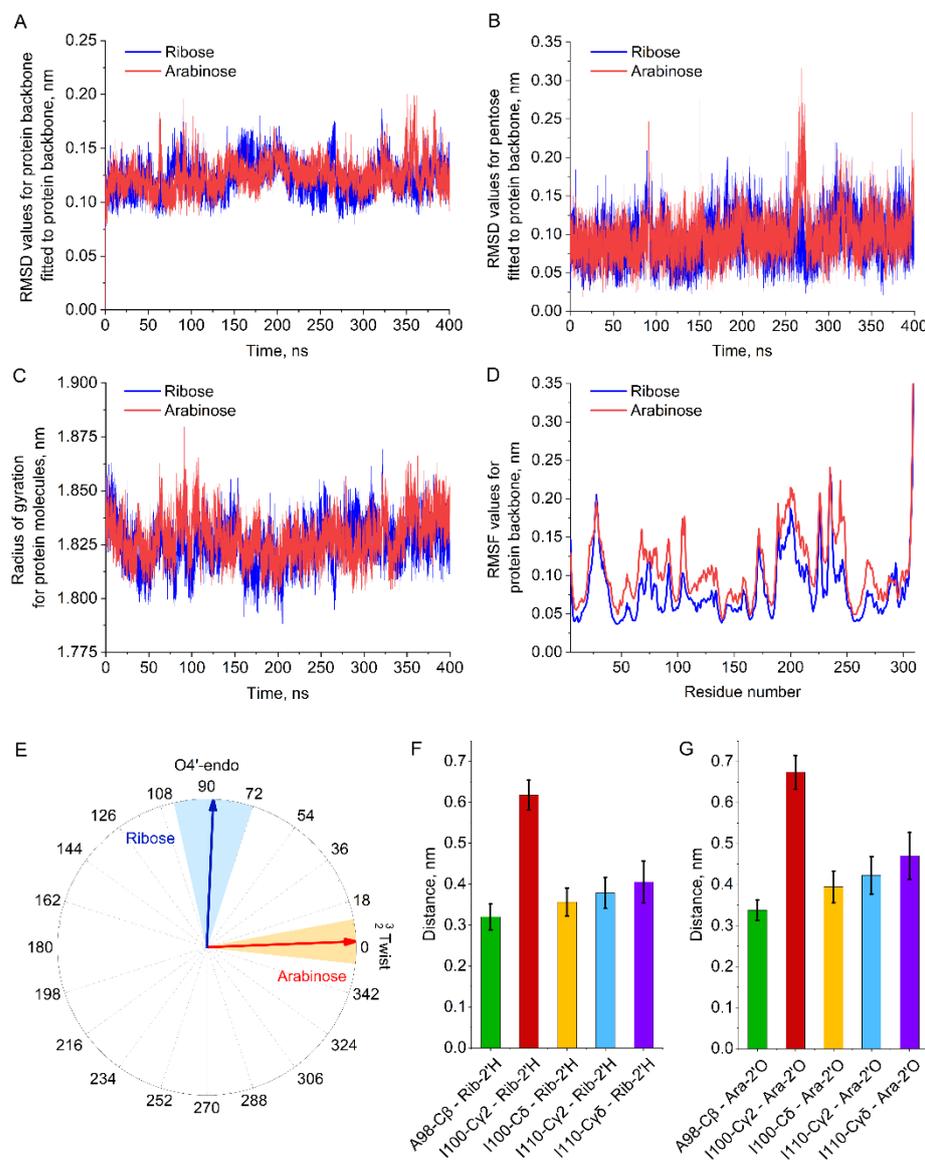


Figure 4. (A)—RMSD values for protein backbone fitted to protein backbone (B)—RMSD values for pentose molecule fitted to protein backbone (C)—radius of gyration for the protein molecule (D)—RMSF profile for the protein backbone (E)—pseudorotation coordinates for the furanose ring of ribose/arabinose (mean value and standard deviation) (F)—Distances between 2'-hydrogen atom of ribose and various atoms of *EcoRK* active site (mean values and standard deviation) (G)—Distances between 2'-oxygen atom of arabinose and various atoms of *EcoRK* active site (mean values and standard deviation).

Further comparative analysis of the positions of arabinose and ribose in the active site of *EcoRK* allows us to determine another factor that may play a role in the enzyme's substrate specificity. In Figures 3 and 5, we can observe that the residues Ala98, Ile100 and Ile110 form a hydrophobic interface in the lid domain. In the case of the natural substrate, this hydrophobic interface is exposed to the non-polar 2' and 3' hydrogen atoms of ribose (Figure 3). However, in the case of our compound of interest, the same interface is

exposed to the polar 2'-hydroxyl group of arabinose (Ara-2'OH), potentially resulting in a unfavorable interaction. Therefore, we can hypothesize that one of the reasons behind the low affinity of *EcoRK* towards arabinose is a steric conflict between Ara-2'OH and the hydrophobic interface of the lid domain formed by Ala98, Ile100 and Ile110. We have decided to focus our enzyme design project on the solution of this possible issue, aiming both to test our hypothesis experimentally and, hopefully, obtain the desired mutant *EcoRK* with elevated affinity towards arabinose.

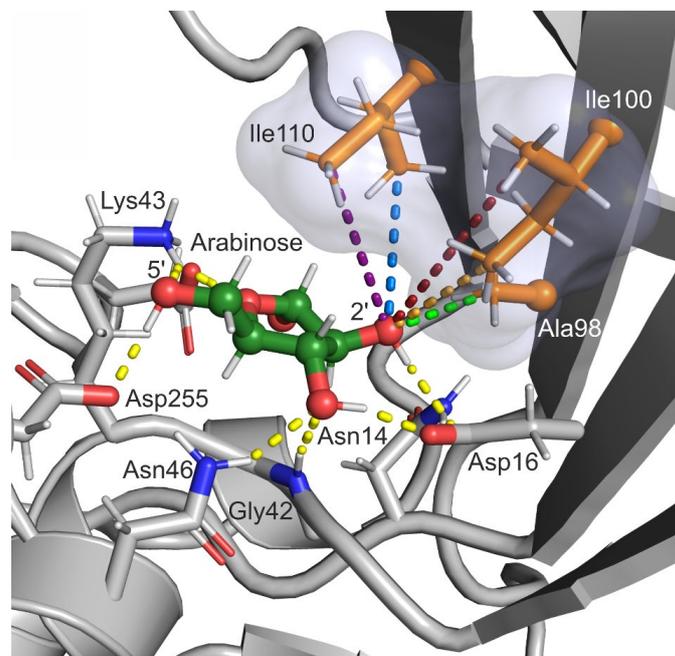


Figure 5. The position of arabinose in the active site of *EcoRK* (structure obtained through cluster analysis of 400-ns long MD trajectory). The semi-transparent surface represents the hydrophobic interface in the lid domain, formed by Ala98, Ile100 and Ile110. Residues 248–251 and 285–295 are hidden. Carbon atoms of arabinose are colored green, carbon atoms of residues Ala98, Ile100 and Ile110—orange, oxygen atoms—red, nitrogen atoms—blue. Yellow dashed lines represent polar contacts; other colors distances measured throughout the MD simulation and represented in Figure 4F,G. Visualization was performed in PyMOL 2.0 (Schrödinger, LLC).

Throughout the MD simulations, the atoms of the three mentioned lid-domain residues remained at the same distance from both 2'-oxygen atom of arabinose and 2'-hydrogen atom of ribose (Figure 4F,G). The closest atom to the 2'-oxygen atom of arabinose was the β -atom of Ala98, followed by the δ -atom of Ile100. Therefore, the two most promising mutations are A98G and I100V. There might be a possibility that these substitutions may act in a synergetic manner, while having no effect individually, making the double A98G/I100V mutant also potentially feasible. In order to study the role of the third mentioned residue, the I110V is also worthy of experimental investigation. Further reduction in the Ile sidechain (mutations to Ala or Gly) may not be feasible before we test the less disruptive substitutions.

To conclude, investigation of the *EcoRK* active site *in silico* has allowed us to hypothesize the reasons behind low affinity of this enzyme towards arabinose and select three hot spots for enzyme design with four mutations worthy of experimental study: A98G, I100V, I110V and A98G/I100V.

2.2. Enzyme Assay, Kinetic Studies and Cascade Synthesis

Having selected substitutions feasible for investigation *in vitro*, we introduced these mutations in the gene encoding the wild-type *EcoRK*, created efficient producer strains and produced and purified the recombinant mutant enzymes (along with the wild type) to a purity of $85 \pm 5\%$ (Figure S1). We then performed the enzyme assay for the wild type

and mutant *EcoRK* determining the ribose and arabinose phosphorylation activity through the detection of conversion of ATP to ADP via HPLC. For the natural substrate ribose, an enzyme assay was performed at an equimolar concentration (2 mM ribose, 2 mM ATP). Since the native *EcoRK* affinity towards arabinose is too low to perform enzyme assay at equimolar concentration, we used a 25-fold excess of arabinose (50 mM arabinose, 2 mM ATP). This concentration is close to the one we have used in our previous study, dealing with the preparative cascade synthesis of arabinosides using wild type enzymes [4]. The obtained results are presented in Figure 6.

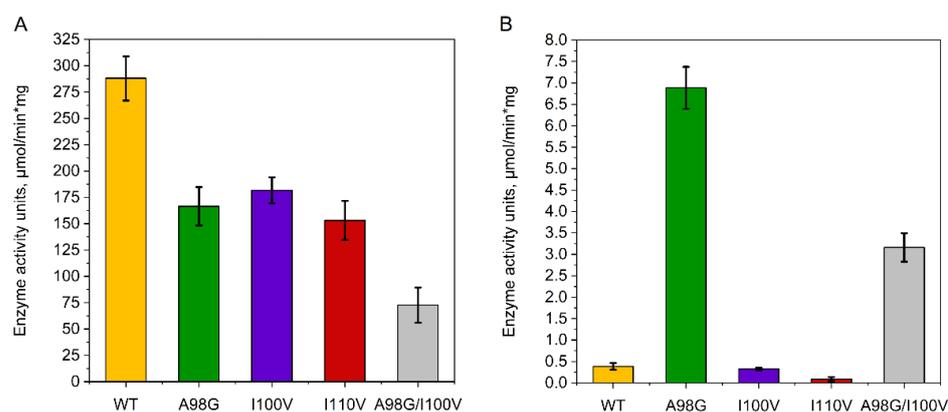


Figure 6. Enzyme assay of wild type *EcoRK* and various mutants for (A) D-ribose phosphorylation (2 mM) (B) D-arabinose phosphorylation (50 mM). Data represented in mean values and standard deviation (in three repeats).

The enzyme assay revealed that all of the investigated mutations decreased *EcoRK* activity towards the natural substrate ribose, with the double mutation demonstrating the most pronounced effect. This result confirms our hypothesis that Ala98, Ile100 and Ile110 residues form a hydrophobic interface that plays a role in ribose binding. Most importantly, the A98G substitution resulted in an increase in *EcoRK* activity towards arabinose by an order of magnitude. The double A98G/I100V mutation has also resulted in a significant positive effect, although the observed increase was less than that of the single A98G mutation. Meanwhile, I100V did not have any significant effect and I110V resulted in a fivefold decrease in enzyme activity. To conclude, we have determined that two *EcoRK* mutants display elevated activity towards D-arabinose, and therefore proceeded to the comparative study of enzyme kinetics.

We have determined the kinetic parameters for the enzymatic activity of the native enzyme, the single A98G and the double A98G/I100V mutants towards D-ribose, D-arabinose and ATP. The results are presented Table 1. Michaelis–Menten plots are presented in Figure S2.

Kinetic studies of the wild type enzyme have shown that arabinose has a two orders of magnitude lower k_{cat} and three orders of magnitude higher K_{M} compared to the natural substrate. Comparison of kinetic parameters illustrates that the A98G mutation caused opposite effects for *EcoRK* activity towards ribose and arabinose. In the case of ribose, this substitution resulted in a twofold increase in K_{M} and decrease in k_{cat} , highlighting the importance of Ala98 for both natural substrate binding and catalysis. On the other hand, for our substrate of interest, D-arabinose, we observed a 15-fold increase in k_{cat} and 1.5-fold decrease in K_{M} . The second I100V mutation has caused a slight decrease in k_{cat} and a twofold increase in K_{M} for arabinose phosphorylation compared to the A98G mutant. In terms of activity towards ATP, each of the two subsequent mutations caused a slight increase in K_{M} .

To conclude, the single A98G mutant turned out to be the most active towards arabinose. With the achieved 15-fold increase in k_{cat} towards arabinose, our enzyme design project turned out to be fully successful. To illustrate the practical applicability of our results, we carried out a comparative study of the performance of the wild

type *EcoRK* and A98G mutant in the cascade synthesis of arabinosides. For this purpose, we chose 2-chloroadenine arabinoside as the test subject, which is a compound closely related to the FDA-approved drugs cladribine (deoxyriboside) and clofarabine (2-deoxy-2-fluoroarabinoside) [24,25]. We have previously reported the cascade synthesis of 2-chloroadenine arabinoside with wild type *EcoRK*, phosphopentomutase and purine nucleoside phosphorylase from *E. coli* (*EcoPPM* and *EcoPNP*, respectively) [4]. We have performed the synthesis of 2-chloroadenine arabinoside using wild type and A98G mutant *EcoRK* in the same reaction conditions, although with a 10-fold reduction in the quantity of the mutant enzyme.

Table 1. Kinetic parameters for the enzymatic activity of wild type *EcoRK*, A98G and A98G/I100V mutants towards D-ribose, D-arabinose and ATP (determination of activity was performed in three repeats).

	V_{max} , $\mu\text{mol}/\text{min}\cdot\text{mg}$	k_{cat} , s^{-1}	K_M , mM
D-ribose:			
Wild type	340 ± 50	180 ± 30	0.15 ± 0.02
A98G	180 ± 20	97 ± 11	0.28 ± 0.03
A98G/I100V	130 ± 30	70 ± 16	0.24 ± 0.05
D-arabinose:			
Wild type	3.6 ± 0.5	1.9 ± 0.3	470 ± 60
A98G	54 ± 5	29 ± 3	310 ± 30
A98G/I100V	33 ± 5	17 ± 3	730 ± 120
ATP:			
Wild type	350 ± 40	185 ± 21	0.43 ± 0.05
A98G	160 ± 10	85 ± 5	0.58 ± 0.04
A98G/I100V	120 ± 10	63 ± 5	0.75 ± 0.09

As a result, the synthesis utilizing 1 μg of A98G mutant was as efficient as that with 10 μg of the wild type enzyme (Figure 7). In both cases, the conversion of 2-chloroadenine to the desired arabinoside product peaked at 60% after 1.5 h of incubation. Meanwhile, synthesis with only 1 μg of native *EcoRK* was highly inefficient: the conversion did not exceed 9% after 3 h. Therefore, we have proven that our engineered A98G mutant allows us to perform the cascade synthesis of arabinosides with an order of magnitude less enzyme quantity with no loss of synthesis efficiency.

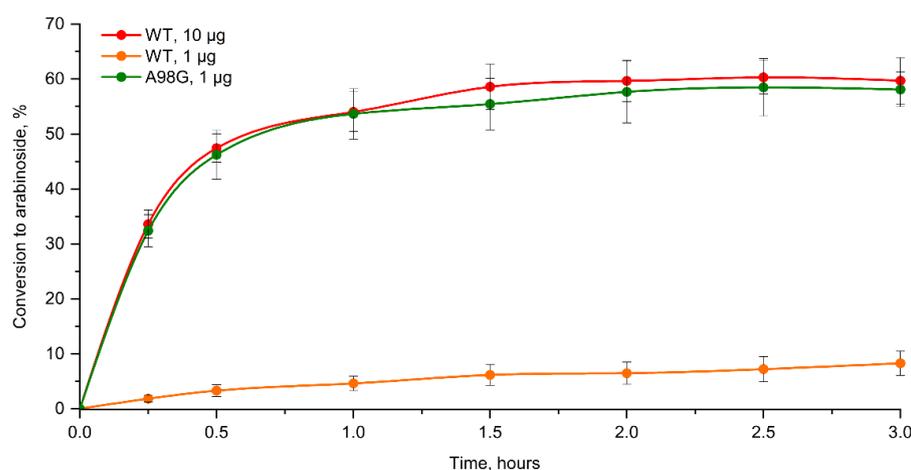


Figure 7. Accumulation of 2-chloroadenine arabinoside during cascade synthesis with the wild type *EcoRK* and the A98G mutant. The data are presented in mean values with standard deviation. Data represented in mean values and standard deviation (in three repeats).

3. Discussion

Our enzyme design project was dedicated to engineering a mutant form of *EcoRK* with elevated activity towards D-arabinose. Molecular dynamics simulations have allowed us to study the binding of arabinose by *EcoRK* in silico and hypothesize the reasons behind the low enzyme activity towards this unnatural substrate. Guided by our hypothesis, we have selected four mutations worthy of investigation in vitro: A98G, I100V, I110V and A98G/I100V. Among these mutations, A98G turned out to be the most efficient, resulting in an order of magnitude 15-fold increase in k_{cat} and 1.5-fold decrease in K_M . Therefore, the target of our research was successfully achieved.

The order of magnitude increase in enzyme activity allows us to lower the amount of recombinant enzyme used in cascade synthesis, which is a major factor limiting the feasibility of the synthesis of arabinosides using wild type enzymes. We have also achieved a 1.5-fold decrease in K_M , which can allow us to perform cascade synthesis using a lower concentration of arabinose. While the K_M value of arabinose phosphorylation by *EcoRK* A98G remains very high, the need for an excess of this substrate is not a critical issue, since D-arabinose is a commodity reagent.

We have illustrated the practical benefits of the A98G mutant through the cascade synthesis of 2-chloroadenine arabinoside as a test subject. Our previous research has shown that state-of-art wild type *EcoRK* requires large amounts of the recombinant enzyme for the efficient synthesis of arabinosides, making the feasibility of the cascade synthesis of these compounds highly questionable. However, the result of this research project, namely the A98G mutant, allows us to perform the cascade synthesis of 2-chloroadenine arabinoside with an order of magnitude less of the enzyme without a loss in substrate conversion. Therefore, with our engineered mutant, the narrow substrate specificity of ribokinase is no longer such a critical issue for the cascade synthesis of arabinosides. Our engineered *EcoRK* A98G mutant can be a basis for future research, dedicated to the development of efficient approaches of cascade synthesis of medically important cancer suppressing and antiviral nucleosides. The narrow carbohydrate substrate specificity of other enzymes of this cascade (namely PPM and PNP) remains a challenge that we are going to address in our future enzyme design projects.

Another aspect of our research that is worthy of discussion are the implications of our results for knowledge concerning the *EcoRK* catalytic mechanism, and the relationship between the enzymes structure and substrate specificity.

Both enzyme assay and kinetic studies of wild type and mutant *EcoRK* activity towards ribose confirm that the Ala98, Ile100 and Ile110 play an important role in the natural reaction. Molecular dynamics simulations have shown that atoms of these three residues are exposed to the non-polar 2'-hydrogen atoms of ribose (Figures 3 and 4F). In context of this structural data, the negative effect of Ala→Gly and Ile→Val substitutions on *EcoRK* activity towards ribose can be explained by the disruption of an important hydrophobic interaction through a reduction in the sidechain size of Ala98, Ile100 and Ile110.

The study of arabinose binding in silico has revealed that this substrate assumes the 3_2 Twist conformation in the active site of *EcoRK*. The 2'-hydroxyl group of arabinose is exposed to the three mentioned lid-domain residues that form a hydrophobic interface. We have observed that throughout the MD simulation the β -carbon atom of Ala98 was the closest to Ara-2'OH. This fact can explain why A98G has caused an order of magnitude increase in k_{cat} towards arabinose, while I100V and I110V did not display any positive effect. The double A98G/I100V mutant had a slightly lower k_{cat} and a twofold higher K_M compared to the single A98G mutant, which further shows that the I100V had no positive effect on arabinose binding.

Since we have measured pentose phosphorylation through the detection of the conversion of ATP to ADP, we also had to take into account the possibility that the effects we have observed may also be caused by the influence of the investigated mutations on *EcoRK* activity towards ATP. Kinetic studies have shown that the A98G and A98G/I100V mutations have increased the K_M for ATP. Therefore, the observed positive effects of A98G

and A98G/I100V mutations on this activity towards arabinose cannot be explained by a change in affinity of *EcoRK* towards the phosphate donor.

To conclude, we have successfully achieved all of the goals of our enzyme-design project. We have developed a mutant form of ribokinase (*EcoRK* A98G) with a 15-fold increase in k_{cat} and twofold decrease in K_{M} towards D-arabinose. We have shown that this mutant allows the cascade synthesis of arabinosides with an order of magnitude less enzyme with no loss in synthesis efficiency. Our mutational studies provide insight into the roles of three hydrophobic residues of the lid domain (Ala98, Ile100 and Ile110) in the substrate specificity of *EcoRK*, and, in particular, explain one of the reasons behind the enzyme's low affinity towards arabinose.

4. Materials and Methods

4.1. Computational Approaches

The analysis, visualization and editing of molecular models were performed in PyMOL 2.0 (Schrödinger, LLC, New York, NY, USA). Simulations of molecular dynamics were carried out using the Gromacs 2021 package modified with PLUMED v2.8.0 plugin [26,27]. The parameters of the molecular models were described using Amber ff99SB-ILDN forcefield [28]. Ligand parameters were obtained through the Antechamber package (AmberTools2022) using the bcc charge calculation method [29].

The structure of the complex between *EcoRK* and its natural substrate ribose was fetched from the Protein Data Bank (PDB 1RKD). The inversion of the chirality of the C2' atom of ribose in this structure via molecular editing was performed to obtain the model of the enzyme in a complex with arabinose (initial approximation). Both of the protein-ligand complexes were placed into the centers of rhombic dodecahedral periodic cells (1000 nm³ in volume, so that any protein atom was at least 1 nm away from the cell edge), filled with tip3p water-molecule models. Systems were neutralized with KCl ions to a concentration of 150 mM.

The obtained systems were subjected to energy minimization via the steepest descent method until no atom was subject to a force larger than 1000 kJ/(M·nm⁻²). The next step was a 0.5-ns long NVT equilibration for 333,15 K using a modified Berendsen (V-rescale) thermostat, followed by 0.5-ns long NPT equilibration for 1 bar using the Parrinello–Rahman barostat [30,31].

Productive molecular dynamics simulations were carried out in an isothermal-isobaric ensemble with the same thermostat and barostat as used during equilibration. Integration with a time step of 2 fs was carried out using the Leap-Frog algorithm [32]. The non-covalent interactions were only taken into account for atoms within 1.4 nm of each other. We used the smooth particle mesh Ewald (PME) summation method with cubic interpolation and grid spacing in Fourier to account for the long-range electrostatic interactions [33]. The bonds were constrained using the LINCS algorithm [34].

We performed productive MD simulations, resulting in two 400-ns long trajectories. These trajectories were modified via the *gmx trjconv* command with the *-pbc* flag so that no molecule was split due to periodic boundary conditions. RMSD, RMSF and gyration radius were calculated using the *gmx rms*, *gmx rmsf* and *gmx gyrate* commands, respectively. Distance between atoms and sugar pseudorotation coordinates were calculated via the PLUMED plugin.

4.2. Genetic Engineering and Producer Strain Cultivation

We have used the plasmid vector from our previous research (Chuvikovsky et al., 2006), containing the *rbsK* gene, encoding ribokinase from *E. coli* (Uniprot P0A9J6), [15]. Mutations in the *rbsK* gene were performed using the Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific, Waltham, MA, USA). The obtained expression vectors (along with the one containing the native gene) were transformed into competent *E. coli* ER2566 (New England Biolabs, Ipswich, MA, USA) strain cells.

The resulting producer strains were cultivated in the Luria–Bertani medium (per 1 L: 10 g. tryptone, 5 g. yeast extract 10 g. NaCl) with 100 µg/mL ampicillin. Cell culture was cultivated until reaching $A_{595} = 0.8$, followed by supplementation with 0.4 mM IPTG (isopropyl β-D-1-thiogalactopyranoside). After induction by IPTG, the cells were cultivated for 4 h at 37 °C, resulting in recombinant enzyme in soluble form.

4.3. Purification of EcoRK

The cells were resuspended in 0.1 M Tris-HCl pH 8, 1 mM PMSF (phenylmethylsulfonyl fluoride) buffer (1:10 *w/v*) and homogenized to ultrasonic disintegration, followed by centrifugation. Cell supernatant was diluted with three volumes of distilled water and applied onto the XK16/10 (GE Healthcare, Chicago, IL, USA) column, packed with 5 mL Q Sepharose High Performance sorbent, pre-equilibrated with 25 mM Tris-HCl and 5 mM EDTA buffer. After washing the column, the recombinant enzyme was eluted with a linear 0–0.5 M NaCl gradient.

The fractions containing the desired recombinant protein were pooled, concentrated to 10 ± 2 mg/mL on the YM-30 membrane (Millipore, Burlington, MA, USA) and applied onto the Superdex 200 column 16×70 mm (Amersham Bioscience, Amersham, UK), pre-equilibrated with the final buffer: 50 mM Tris-HCl pH8, 50 mM KCl, 5 mM MgCl₂, 5% glycerol, 0.04% NaN₃. The eluted purified protein was used for further experiments. We used the Bradford method to determine protein concentration and polyacrylamide gel electrophoresis under denaturing conditions for the determination of protein purity [35,36].

4.4. Enzyme Assay and Determination of Kinetic Parameters

The 100 µL reaction mix for the enzyme assay included: 2 mM D-ribose or 50 mM D-arabinose, 2 mM ATP, 50 mM Tris-HCl pH8, 50 mM KCl, 5 mM MgCl₂, 1 mM KH₂PO₄. The samples were pre-heated, followed by the addition of the investigated enzyme in a quantity of 1.5×10^{-2} µg for ribose and 1 µg for arabinose. Samples supplemented with the studied enzyme were incubated for four minutes at 37 °C, after which an enzyme assay was performed via analytical HPLC (in three repeats).

The conversion of the substrate into the product was observed through the detection of the ATP and ADP via HPLC. Analytical HPLC was performed on the Waters system (pump 1525, detector 2489, Breeze Software, Milford, NH, USA) with the following method: Supelcosil LC-18-T, 5 µm, 150×4.6 mm column, isocratic elution with 100 mM KH₂PO₄ pH6, flow rate 0.4 mL/min, detection at 254 nm. One unit of EcoRK activity was defined as 1 µmol of pentose substrate converted into product per minute per milligram of recombinant protein.

The reaction mix (100 µL) for the determination of the kinetic parameters of ribose phosphorylation included: 0.06–6 mM D-ribose, 2 mM ATP, 50 mM Tris-HCl pH8, 50 mM KCl, 5 mM MgCl₂, 1 mM KH₂PO₄ and 1.5×10^{-2} µg of the investigated enzyme. The reaction mix (100 µL) for arabinose included: 50–1600 mM D-arabinose, 2 mM ATP, 50 mM Tris-HCl pH8, 50 mM KCl, 5 mM MgCl₂, 1 mM KH₂PO₄ and 1 µg of wild type EcoRK or 0.15 µg of the mutant enzyme. The reaction mix (100 µL) for the determination of kinetic parameters with ATP as the variable substrate included: 0.1–5 mM ATP, 2 mM D-ribose, 50 mM Tris-HCl pH8, 50 mM KCl, 5 mM MgCl₂, 1 mM KH₂PO₄ and 1.5×10^{-2} µg of the investigated enzyme. Enzyme assay for the varying substrate concentrations were performed as described above. We determined the kinetic parameters through nonlinear regression analysis using the SciDAVis v2.3.0 program. Catalytic constants were calculated for one enzyme subunit (32.3 kDa according to amino acid sequence). The enzyme kinetics model was fitted to the Michaelis–Menten equation.

4.5. Cascade Synthesis of 2-Chloroadenine Arabinoside

The reaction mix (500 µL) for the cascade synthesis of 2-chloroadenine arabinoside included: 60 mM D-arabinose, 0.34 mM 2-chloroadenine; 2.5 mM ATP; 50 mM Tris-HCl

pH8, 2.5 mM MnCl₂, 50 mM KCl, 2 mM KH₂PO₄, ribokinase from *E. coli* (1 or 10 µg of the wild type enzyme, or 1 µg of the A98G mutant), 200 µg of phosphopentomutase from *E. coli* and 40 µg of purine nucleoside phosphorylase from *E. coli*. Recombinant phosphopentomutase from *E. coli* and purine nucleoside phosphorylase from *E. coli* were obtained in sufficient quantities according to the methods provided in their respective publications [10,37].

The reaction mixes with the wild type and A98G mutant *EcoRK* were incubated at 50 °C for 3 h. The conversion of 2-chloroadenine into 2-chloroadenine arabinoside was detected through analytical HPLC on the Waters system (Waters 1525, Waters 2489, Breeze) in three repeats with the following method: column YMC Triart-C18, 50 × 3.0 mm, 3 µm, eluent 0.1% aqueous trifluoroacetic acid and 7% acetonitrile, detection at 254 nm, flow rate 0.4 mL/min.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232012540/s1>.

Author Contributions: Conceptualization, E.A.Z., I.D.K. and R.S.E.; methodology, E.A.Z., I.V.F. and M.A.K.; software, E.A.Z. and V.I.T.; validation, I.V.F. and M.A.K.; formal analysis, I.V.F., M.Y.B. and B.Z.E.; investigation, E.A.Z., I.V.F., M.A.K., D.O.Y., Y.A.A., D.D.L., M.Y.B., B.Z.E. and V.I.T.; resources, I.V.F. and M.A.K.; data curation, I.D.K. and R.S.E.; writing—original draft preparation, E.A.Z. and I.V.F.; writing—review and editing, E.A.Z., I.D.K. and R.S.E.; visualization, E.A.Z.; supervision, I.D.K. and R.S.E.; project administration, I.D.K. and R.S.E.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation (Project No. 21-13-00429).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Mikhailopulo, I.A.; Miroshnikov, A.I. New trends in nucleoside biotechnology. *Acta Nat.* **2010**, *2*, 36–59. [CrossRef]
2. Jordheim, L.P.; Durantel, D.; Zoulim, F.; Dumontet, C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat. Rev. Drug. Discov.* **2013**, *12*, 447–464. [CrossRef] [PubMed]
3. Roy, B.; Depaix, A.; Périgaud, C.; Peyrottes, S. Recent Trends in Nucleotide Synthesis. *Chem. Rev.* **2016**, *116*, 7854–7897. [CrossRef] [PubMed]
4. Fateev, I.V.; Kostromina, M.A.; Abramchik, Y.A.; Eletskaia, B.Z.; Mikheeva, O.O.; Lukoshin, D.D.; Zayats, E.A.; Berzina, M.Y.; Dorofeeva, E.V.; Paramonov, A.S.; et al. Multi-Enzymatic Cascades in the Synthesis of Modified Nucleosides: Comparison of the Thermophilic and Mesophilic Pathways. *Biomolecules* **2021**, *11*, 586. [CrossRef]
5. Scism, R.A.; Bachmann, B.O. Five-component cascade synthesis of nucleotide analogues in an engineered self-immobilized enzyme aggregate. *ChemBioChem* **2010**, *11*, 67–70. [CrossRef]
6. Esipov, R.S.; Abramchik, Y.A.; Fateev, I.V.; Konstantinova, I.D.; Kostromina, M.A.; Muravyova, T.I.; Artemova, K.G.; Miroshnikov, A.I. A Cascade of Thermophilic Enzymes As an Approach to the Synthesis of Modified Nucleotides. *Acta Nat.* **2016**, *8*, 82–90. [CrossRef]
7. Yehia, H.; Kamel, S.; Paulick, K.; Wagner, A.; Neubauer, P. Substrate spectra of nucleoside phosphorylases and their potential in the production of pharmaceutically active compounds. *Curr. Pharm. Des.* **2017**, *23*, 6913–6935. [CrossRef]
8. Almendros, M.; Berenguer, J.; Sinisterra, J.V. *Thermus thermophilus* nucleoside phosphorylases active in the synthesis of nucleoside analogues. *Appl. Environ. Microbiol.* **2012**, *78*, 3128–3135. [CrossRef]
9. Fateev, I.V.; Kharitonova, M.I.; Antonov, K.V.; Konstantinova, I.D.; Stepanenko, V.N.; Esipov, R.S.; Seela, F.; Temburnikar, K.W.; Seley-Radtke, K.L.; Stepchenko, V.A.; et al. Recognition of Artificial Nucleobases by *E. coli* Purine Nucleoside Phosphorylase versus its Ser90Ala Mutant in the Synthesis of Base-Modified Nucleosides. *Chemistry* **2015**, *21*, 13401–13419. [CrossRef]
10. Miroshnikov, A.I.; Esipov, R.S.; Muravyova, T.I.; Konstantinova, I.D.; Fateev, I.V.; Mikhailopulo, I.A. A New Strategy for the Synthesis of Nucleosides: One-Pot Enzymatic Transformation of D-Pentoses into Nucleosides. *Open. Conf. Proc. J.* **2010**, *1*, 98–102. [CrossRef]

11. Sigrell, J.A.; Cameron, A.D.; Jones, T.A.; Mowbray, S.L. Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: Insights into a new family of kinase structures. *Structure* **1998**, *6*, 183–193. [[CrossRef](#)]
12. Sigrell, J.A.; Cameron, A.D.; Mowbray, S.L. Induced fit on sugar binding activates ribokinase. *J. Mol. Biol.* **1999**, *290*, 1009–1018. [[CrossRef](#)] [[PubMed](#)]
13. Timofeev, V.I.; Abramchik, Y.A.; Muravyova, T.I.; Zhukhlistova, N.E.; Esipov, R.S.; Kuranova, I.P. Three-Dimensional Structure of Recombinant Thermophilic Ribokinase from *Thermus* species 2.9 in Complex with Adenosine Diphosphate. *Crystallogr. Rep.* **2021**, *66*, 769–776. [[CrossRef](#)]
14. Paul, R.; Patra, M.D.; Sen, U. Crystal structure of apo and ligand bound vibrio cholerae ribokinase (Vc-RK): Role of monovalent cation induced activation and structural flexibility in sugar phosphorylation. *Adv. Exp. Med. Biol.* **2015**, *842*, 293–307. [[CrossRef](#)] [[PubMed](#)]
15. Chuvikovskiy, D.V.; Esipov, R.S.; Skoblov, Y.S.; Chupova, L.A.; Muravyova, T.I.; Miroshnikov, A.I.; Lapinjoki, S.; Mikhailopulo, I.A. Ribokinase from *E. coli*: Expression, purification, and substrate specificity. *Bioorg. Med. Chem.* **2006**, *14*, 6327–6332. [[CrossRef](#)]
16. Esipov, R.S.; Abramchik, Y.A.; Fateev, I.V.; Muravyova, T.I.; Skoblov, Y.S.; Kostromina, M.A.; Miroshnikov, A.I. Preparation and study of the substrate specificity of thermophilic ribokinase from *Thermus* sp. 2.9. *Russ. J. Biopharm.* **2016**, *8*, 3–12.
17. Lukenbill, J.; Kalaycio, M. Fludarabine: A review of the clear benefits and potential harms. *Leuk. Res.* **2013**, *37*, 986–994. [[CrossRef](#)]
18. Kisor, D.F. Nelarabine: A nucleoside analog with efficacy in T-cell and other leukemias. *Ann. Pharmacother.* **2005**, *39*, 1056–1063. [[CrossRef](#)]
19. Murphy, T.; Yee, K. Cytarabine and daunorubicin for the treatment of acute myeloid leukemia. *Expert. Opin. Pharmacother.* **2017**, *18*, 1765–1780. [[CrossRef](#)]
20. Whitley, R.; Alford, C.; Hess, F.; Buchanan, R. Vidarabine: A preliminary review of its pharmacological properties and therapeutic use. *Drugs* **1980**, *20*, 267–282. [[CrossRef](#)]
21. Park, J.; van Koeverden, P.; Singh, B.; Gupta, R.S. Identification and characterization of human ribokinase and comparison of its properties with *E. coli* ribokinase and human adenosine kinase. *FEBS Lett.* **2007**, *581*, 3211–3216. [[CrossRef](#)] [[PubMed](#)]
22. Fateev, I.V.; Antonov, K.V.; Konstantinova, I.D.; Muravyova, T.I.; Seela, F.; Esipov, R.S.; Miroshnikov, A.I.; Mikhailopulo, I.A. The chemoenzymatic synthesis of clofarabine and related 2'-deoxyfluoroarabinosyl nucleosides: The electronic and stereochemical factors determining substrate recognition by *E. coli* nucleoside phosphorylases. *Beilstein. J. Org. Chem.* **2014**, *10*, 1657–1669. [[CrossRef](#)] [[PubMed](#)]
23. Kang, P.A.; Oh, J.; Lee, H.; Witte, C.P.; Rhee, S. Crystal structure and mutational analyses of ribokinase from *Arabidopsis thaliana*. *J. Struct. Biol.* **2019**, *206*, 110–118. [[CrossRef](#)] [[PubMed](#)]
24. Liliemark, J. The clinical pharmacokinetics of cladribine. *Clin. Pharmacokinet.* **1997**, *32*, 120–131. [[CrossRef](#)]
25. Pui, C.H.; Jeha, S. Clofarabine. *Nat. Rev. Drug. Discov.* **2005**, *4*, 369–370. [[CrossRef](#)]
26. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [[CrossRef](#)] [[PubMed](#)]
27. PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670–673. [[CrossRef](#)]
28. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958. [[CrossRef](#)]
29. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260. [[CrossRef](#)]
30. Berendsen, H.J.C.; Postma, J.P.M.; Van Gunsteren, W.F.; Dinola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690. [[CrossRef](#)]
31. Parrinello, M.; Rahman, A. Strain fluctuations and elastic constants. *J. Chem. Phys.* **1982**, *76*, 2662–2666. [[CrossRef](#)]
32. Van Gunsteren, W.F.; Berendsen, H.J.C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1*, 173–185. [[CrossRef](#)]
33. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8592. [[CrossRef](#)]
34. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472. [[CrossRef](#)]
35. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [[CrossRef](#)]
36. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **1970**, *227*, 680–685. [[CrossRef](#)]
37. Esipov, R.S.; Gurevich, A.I.; Chuvikovskiy, D.V.; Chupova, L.A.; Muravyova, T.I.; Miroshnikov, A.I. Overexpression of *Escherichia coli* genes encoding nucleoside phosphorylases in the pET/BI21(DE3) system yields active recombinant enzymes. *Protein Expr. Purif.* **2002**, *24*, 56–60. [[CrossRef](#)]