Thermostable Nucleoside Phosphorylases as Biocatalysts for the Synthesis of Purine Nucleoside Analogues

Characterisation, immobilization and synthesis

Institut für Biotechnologie
Technische Universität Berlin
Bioverfahrenstechnik

Xinrui Zhou, Berlin 2014

Thermostable Nucleoside Phosphorylases as Biocatalysts for the Synthesis of Purine Nucleoside Analogues

Characterisation, immobilization and synthesis

vorgelegt von
M. Sc. Xinrui Zhou
aus Kunming (China)

von der Fakultät III – Prozesswissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Dr. rer.nat. –
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender	Prof. Dr. Roland Lauster
Gutachter	Prof. Dr. Peter Neubauer
Gutachter	Prof. Dr. Igor A. Mikhailopulo
Gutachter	Dr. habil. Roland Wohlgemuth
Gutachter	Prof. Dr. Vera Meyer

Tag der wissenschaftlichen Aussprache: 14.02.2014

Berlin 2014
D83

The present work was performed from January 2011 to December 2013 in the Laboratory of Bioprocess Engineering at the Department of Biotechnology and at the Department of Chemistry, Technische Universität Berlin under the supervision of Prof. Dr. Peter Neubauer (Technische Universität Berlin) and Prof. Dr. Mikhailopulo (National Academy of Sciences of Belarus).

When I consider Your heavens, the work of Your fingers,

The moon and the stars, which You have ordained,

What is man that You are mindful of him, and the son of man that You visit him?

I will praise You, O Lord, with my whole heart;

I will tell of all Your marvellous works.

I will be glad and rejoice in You;

I will sing praise to Your name, O Most High.

Bible [Psalm 8:3-4, 9:1-2]

Abstract

Nucleosides represent one class of fundamental building blocks of life systems. Their analogues are extensively used as therapeutic agents for cancer and viral diseases; as precursors of oligonucleotides for therapeutic or diagnostic use; and as molecular tools in research. However, many biologically active nucleosides are difficult to synthesize chemically, which hinders biological trials and studies as well as the application of these compounds. The chemo-enzymatic synthesis in comparison to the pure chemical synthesis offers great benefits as a simple, efficient and "green" technology in view of simplicity, costs and yield. The focus of this study was on the application of novel nucleoside phosphorylases (NPs) for the biocatalytic synthesis of nucleoside analogues. The applied enzymes originate from thermo-philic microorganisms (*Deinococcus geothermalis, Geobacillus thermoglucosidasius, Thermus thermphilus,* and A*eropyrum pernix*). Five thermostable NPs (TtPyNP, GtPyNP, DgPNP, GtPNP and ApMTAP) were successfully recombinantly expressed in soluble and active form in the Gram-negative bacterium *Escherichia coli*, and were characterised with respect to temperature optimum, stability, substrate specificity, kinetic behaviour, pH optimum and in view of their similarity by sequence alignments. The results demonstrate their promising biocatalytic properties, especially for TtPyNP and GtPNP, which displayed particularly high catalytic activity towards natural substrates (3-6 times higher than the commonly used enzymes from *E. coli*) under the experimental conditions. Furthermore, they showed high promiscuity, i.e. they were able to catalyse a broad range of modified substrates. Both enzymes accept nucleosides modified in the sugar moiety, e.g. 2'- or 3'-NH_2, 2'-F (*ribo*- or *arabino*-) and 2'-OH (*arabino*); GtPNP recognizes 2,6-Cl or F substituted purine. In coupled reactions with TtPyNP and GtPNP as catalysts various modified purine nucleosides were successfully synthesized. For the synthesis of purine nucleoside analogues, TtPyNP and GtPNP were successfully immobilized on magnetic microsphere beads with high residual enzyme activity, high enzyme loading, and further enhanced enzyme stability. The application of the immobilized enzymes in the synthesis of 2,6-dichloropurine riboside and 6-chloro-2-fluoropurine riboside (6C2FP-R) resulted in product yields of 78.5 % and 85.5 % (HPLC), the latter molecule was isolated and purified (>98 %) by silica chromatography (normal-phase column) and the compound's structure was confirmed. These results reveal the great practical potential of the studied biocatalysts. Hence, it is conceivable to produce a number of nucleoside analogues by the described method, which appears more efficient than other synthetic routes described in literature so far.

Zusammenfassung

Nukleoside sind bedeutende Bausteine lebender Systeme. Chemisch modifizierte Nukleoside sind hochwirksame Pharmaka für die Behandlung von Krebs und viralen Erkrankungen. Sie sind wichtige Werkzeuge in der Zellforschung, u.a. sind sie Vorstufen für die Herstellung von Oligonukleotiden für theapeutische und diagnostische Anwendungen. Allerdings ist die chemische Synthese von modifizierten Nukleosiden in den meisten Fällen kompliziert und mit niedrigen Ausbeuten verbunden, wodurch ihr Einsatz limitiert ist. Die chemo-enzymatische Synthese von Nukleosidanaloga bietet im Vergleich zur traditionellen chemischen Synthese als eine einfache und „grüne" Technologie entscheidende Vorteile in Bezug auf Aufwand, Reaktionsausbeuten und Kosten. In der vorliegenden Arbeit liegt der Fokus auf neu isolierten Nukleosidphosphorylasen von verschiedenen thermophilen Mikroorganismen (*Deinococcus geothermalis, Geobacillus thermoglucosidasius, Thermus thermophilus, Aeropyrum pernix)*. Fünf thermostabile Nukleosidphosphorylasen (TtPyNP, GtPyNP, DgPNP, GtPNP and ApMTAP) wurden rekombinant in dem Gram-negativen Bakterium *Escherichia coli* exprimiert, aufgereinigt und umfassend in Bezug auf die Temperatur- und pH Optima, Stabilität, Substratspezifität, ihr kinetisches Verhalten charaktersisiert, sowie ihre phylogenetische Sequenzähnlichkeit analysiert. Diese Charakteristika demonstrieren die vorteilhaften biokatalytischen Eigenschaften dieser Biokatalysatoren, welches besonders für die Enzyme TtPyNP und GtPNP zutrifft. Unter Reaktionsbedinungen, die für die biokatalytische Syntheses interessant sind, zeichnen sich diese Enzyme nicht nur durch eine hohe katalytische Aktivität gegenüber natürlichen Substraten aus, sondern auch durch ein hohes Maß an Promiskuität, d.h. eine große Anzahl modifizierter Verbindungen als Substrate nutzen können. Beide Enzyme akzeptieren Zucker, die z.B. in folgenden Positionen modifiziert sind: 2'- oder 3'-NH_2, 2'-F (*ribo*- oder *arabino*-) und 2'-OH (*arabino*). GtPNP reagiert auch mit 2,6-Cl- oder F-substituierten Purinen. Zusammen können TtPyNP und GtPNP die Synthese verschiedener modifizierter Purinnukleoside katalysieren. Für die Anwendung wurden beide Enzyme auf magnetischen mikrospherischen Trägern immobilisiert. Die Immobilisate zeigten eine hohe Aktivität und eine hohe katalytische Dichte, sowie eine erhöhte Enzymstabilität. Bei der Anwendung dieser immobilisierten Biolkatalysatoren für die Synthese von

2,6-Dichloropurine Ribosid (2,6CP-R) und 6-Chloro-2-Fluoropurine Ribosid (6C2FP-R) wurden Produktausbeuten von 78.5 % bzw. 85.5 % erreicht. 6C2FP-R wurde auf >98% über eine Standard Silizium-Chromatographiesäule aufgereinigt und die Struktur verifiziert. Diese Ergebnisse zeigen das große praktische Potenzial der hier untersuchten Biokatalysatoren. Es ist daher naheliegend, dass einige biologisch wichtige Nucleosid durch das in dieser Arbeit beschriebene Verfahren effizienter als wie bisher in der Literatur beschrieben, hergestellt werden können.

Acknowledgements

I am indebted to so many people for their encouragement and direct or indirect support which was essential for the successful completion of this research work. I am aware of that my words are so limit to express my deep gratitude and respect for all of them.

I would like to express my greatest appreciation to my supervisor Professor Peter Neubauer. He offered me chance to work on my PhD in Germany and gave me this fascinating project. His brilliant ideas, endless energy, and his enthusiasm for biotechnology in both academic and industrial areas, greatly promotes the project progress. His constant encouragement, patient guidance and, especially during my thesis writing his kindness, gentleness and full support have touched me so much.

I would like to express my deepest gratitude to my second supervisor Professor Igor A. Mikhailopulo, although we are separated by great distances, his unreserved long-last support, valuable and constructive suggestions come along with every move of the project. He also gave much attention and time to read, correct and make detailed comments on my reports and manuscripts. I cannot thank you more.

I am grateful to my thesis committee members: Dr. Roland Wohlgemuth from Sigma-Aldrich, Professor Vera Meyer and Professor Roland Lauster from TU Berlin for their precious time and effort. Especially I am thankful to Dr. Roland Wohlgemuth for his great interests in this project and for our fruitful discussions through telephone conferences. His kind donation of the valuable intermediates a-D-Ribose-1-phosphate and 2'-deoxyl-a-D-ribose-1-phosphate allow us to investigate the reverse reactions in more details.

I am very thankful to Kathleen Szeker for working closely with me in the first one and half years on this topic: we planning experiments together, discussing results together and sharing so many exciting moments together (e.g. watching the product peak appearing on the HPLC). I will not forget these happy memories.

I am very appreciative towards my colleagues Jian Li and Jennifer Jaitzig for helping me on the LC/MS measurements of our products in Prof. Roderich Süssmuth group on Saturdays, and towards to the BIG-NSE student Lin-Yu Jiao for our pleasant cooperation on final product purification, and his great help on NMR measurements and mass spectroscopy analysis. My sincere thanks also go to Dr. Nicolas Cruz-Bournazou for his kind help on Design of Experiments and many interesting discussions as well as working with his student Hendrik Glenzer on the enzymatic reaction modelling. I am grateful to Dr. Thomas Böhme and Dr. Bernd Janocha for their contributions on ApMTAP and DgPNP expression and characterisation, also for our enjoyable working time and our friendship. I want to thank Dr. Stefan Junne, Julia Glazyrina and Christian Reitz for their help on our first "fermentation". And thanks to Ursula Pfeiffer for her

interests on this topic and deciding to work for screening new enzymes. Many thanks to our "BioNucleo" team (Michael Raven, Anke Wagner and Maryke Fehlau) for the pleasant team work and inspired discussions.

I am also thankful to Dr. Justin Jordaan from ReSyn Biosciences for the kind donation of the magnetic beads as the immobilization carriers used in this study Moreover, his valuable suggestions, instant answer, as well as carefully revise of the immobilization manuscript, have been very much appreciated. I am also grateful to Dr. Alexander Scholz for the preliminary cooperation and useful advice on immobilization. And many thanks to Professor Alex Azhayev from Metkinen Chemistry for kindly donation of substrates FanaU, aAdo and aIno; to Dr. Birgit Wilding from TU Graz for providing us 7-deazapurine bases to test.

I am deeply thankful for the scholarship funding by the Berlin International Graduate School of Natural Sciences and Engineering (BIG-NSE) of the Cluster of Excellence "Unifying concepts in Catalysis" (UniCat). Many thanks to Dr. Jean-Philippe Lonjaret for his great efforts to bring me here. And the financial support by ESF-Stependium der TU Berlin is gratefully acknowledged.

During the last 3 years, I had many colleagues at BVT and I like to thank all of them for the nice working environment, Biocatalysis-Day (seminar), cake day, cleaning day, translations, German learning, so many helping hands and cheerful hearts. Mirja Krause, Irmgard Maue-Mohn, Dirk Itzeck, and Brigtitte Burckhardt for their steadfast drive towards a cleaner lab, their material and practical support. Herta Klein-Leuendorf for help with bureaucracy. Ping Lu for always waiting for me to go home and nice talk about life and study. Friederike Hillig for sharing rooms with me, her excellent translation service and timely help when I needed (printing my thesis). Julia Glazyrina for her warm embrace and smiling. Erich Kielhorn for his German course and relaxing talk. Dr. Andreas Knepper, Florian Glauche, Eva Brand, Andri hutari, Anika Bockisch, Basant El Kady, Funda Cansu Erterm…for creating a great working atmosphere.

I also would like to give thanks to my Master's supervisors Professor Xiaoda Yang and Professor Qi Wang, for their kind concern and exchanging ideas about science and life.

Finally I want to thank my family members, my dear brothers and sisters in the Gemeinde, for their love, care, encouragement, and prayer for me; and my lovely neighbours who cook food for me during my thesis writing. My deepest appreciation has to go to my parents, for their undeserved love, support and understanding, which lead me going forward.

Original papers

The presented dissertation is based on the following papers:

Paper I:

Szeker, K., Zhou, X., Schwab, T., Casanueva, A., Cowan, D., Mikhailopulo, I. A., Neubauer, P., 2012. Comparative investigations on thermostable pyrimidine nucleoside phosphorylases from *Geobacillus thermoglucosidasius* and *Thermus thermophilus*. Journal of Molecular Catalysis B: Enzymatic. 84, 27-34.

Paper II:

Zhou, X., Szeker, K., Janocha, B., Böhme, T., Albrecht, D., Mikhailopulo, I. A., Neubauer, P., 2013. Recombinant purine nucleoside phosphorylases from thermophiles: preparation, properties and activity towards purine and pyrimidine nucleosides. FEBS Journal. 280, 1475-1490.

Paper III:

Zhou, X., Mikhailopulo, I. A., Neubauer, P., 2014. Immobilization of thermostable nucleoside phosphorylases on MagReSyn™ epoxide microspheres and its application for the synthesis of 2,6-dihalogenated purine nucleosides. (Manuscript to be submitted)

Paper IV:

Zhou, X., Szeker, K., Jiao, L.-Y., Mikhailopulo, I. A., Neubauer, P., 2014. Enzymatic synthesis of 2,6-dihalogenated purine Nucleosides. (Manuscript to be submitted)

Abbreviations

26DCP	2,6-Dichloropurine
26DCP-dR	2,6-Dichloropurine-2′-deoxyriboside
26DCP-R	2,6-Dichloropurine riboside
2CA	2-Chloroadenine
2FA	2-Fluroadenine
6C2FP	6-Chloro-2-fluoropurine
6C2FP-dR	6-Chloro-2-fluoropurine-2′-deoxyriboside
6C2FP-R	6-Chloro-2-fluoropurine riboside
A. pernix or Ap	*Aeropyrum pernix*
aAdo	2′-Amino-2′-deoxyadenosine
aAdo	2′-Amino-2′-deoxyadenosine
ADA	Adenosine deaminase
Ade	Adenine
AdoP	Adenosine phosphorylase
AhPNP	PNP from *Aeromonas hydrophila*
aIno	2′-Amino-2′-deoxyinosine
AMPDA	Adenylate deaminase
AnhU	O^2,2′-Anhydro-1-(β-D-arabinofuranosyl)uracil
ApMTAP	MTAP from *A. pernix*
ApUP	UP from *A. pernix*
AraA	9-(β-D-Arabinofuranosyl)adenine
AraU	1-(β-D-Arabinofuranosyl)uracil
ASOs	Antisense oligonucleotides
aThd	3′-Amino-2′,3′-deoxythymidine
aUrd	2′-Amino-2′-deoxyuridine
AZT	Azidothymidine
B. cereus or Bc	*Bacillus cereus*
B. subtilis or Bs	*Bacillus subtilis*

CAI	Codon adaption index
Cyd	Cytidine
Cyt	Cytosine
D. geothermalis or Dg	*Deinococcus geothermalis*
dAdo	2′-Deoxyadenosine
DAP	2,6-Diaminopurine
dCyd	2′-Deoxycytidine
DERA	deoxyriboaldolase
DgPNP	PNP from D. geothermalis
dGuo	2′-Deoxyguanosine
DHAP	Dihydroxyacetone monophosphate
dIno	2′-Deoxyinosine
DMSO	Dimethyl sulfoxide
dR-1-P	2-deoxyl-a-D-ribose-1-phosphate
DTT	Dithiothreitol
EcPNP	PNP from *E. coli*
EcTP	TP from *E. coli*
EcUP	UP from *E. coli*
FA	2′-Deoxy-2′-fluoroadenosine
FanaA	9-(2-Deoxy-2-fluoro-β-D-arabinofuranosyl)adenine
FanaU	1-(2-Deoxy-2-fluoro-β-D-arabinofuranosyl)uracil
FU	2′-Deoxy-2′-fluorouridine
G. stearothermophilus or Gs	*Geobacillus stearothermophilus*
G. thermoglucosidasius or Gt	*Geobacillus thermoglucosidasius*
Gla-3P	D-glyceraldehyde-3-phosphate
GsPyNP	PyNP from *G. stearothermophilus*
GtPNP	PNP from *G. thermoglucosidasius*
GtPyNP	PyNP from *G. thermoglucosidasius*
Guo	Guanosine

his-tag	Histidine tag
HPLC	High-performance liquid chromatography
IPTG	Isopropyl -β-D-1-thiogalactopyranoside
KP buffer	Potassium phosphate buffer
LB medium	Lysogeny broth medium
MTAP	5'-Methylthioadenosine phosphorylase
NDT	Nucleoside deoxyribosyltransferase
NP	Nucleoside phosphorylase
NP buffer	Sodium phosphate buffer
P. furiosus or Pf	*Pyrococcus furiosus*
PF-1-P	a-D-pentofuranose-1-phosphate
PF-5-P	D-pentose-5-phosphate
PNP	Purine nucleoside phosphorylase
PPM	Phosphopentomutase
PyNP	Pyrimidine nucleoside phosphorylase
R-1-P	a-D-Ribose-1-phosphate
R-5-P	5-phospho-D-ribofuranose
RK	Ribokinase
S. solfataricus or Ss	*Sulfolobus solfataricus*
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel
T. thermophilus or Tt	*Thermus thermophilus*
TB medium	Terrific broth medium
TCA	Trichloroacetic acid
Thd	Thymidine

Preface

The aim of this thesis was to develop an efficient and "green" method – enzymatic synthesis – for the preparation of pharmaceutically or biologically important nucleoside analogues.

In Chapter 1 the most important aspects of enzymatic synthesis of nucleosides are reviewed regards to the sorts of enzymes for the synthesis of nucleosides (Chapter 1.1); the introduction of the focused enzymes in the present study – nucleoside phosphorylases (NPs) – enzyme source, classification and structure (Chapter 1.2); and the modes of application of NPs are also addressed and discussed (Chapter 1.3).

Chapter 1 points out the challenges of using biocatalysts in this area and leads to the present study investigated logically on (i) the preparation and characterization of recombinant NPs from thermophilic microorganisms (Chapter 2), (ii) the study of transglycosylation reactions catalyzed by the NPs (Chapter 3), (iii) immobilization of the NPs and their application for the synthesis of nucleosides (Chapter 4), and (iv) enzymatic synthesis of 2,6-dihalogenated purine nucleosides (Chapter 5).

Chapter 2 covers the main data of published Paper I & Paper II as well as the unpublished results and discussions e.g. the main factors for the expression of thermozymes in *E. coli* (Chapter 2.2); an overview of the substrate specificity of the NPs, pH optimum of the NPs, the enzyme structure-activity relationship based on the multiple protein sequence alignments and the phylogenetic analysis of the interested NPs (Chapter 2.3).

Chapter 3 focuses on the analysis of the processes of the transglycosylation reactions by the characterised free NPs. According to the reaction equilibrium states, the reactions were divided into three types. For each type, the overall (one-pot transglycosylation) and sub-step (phosphorolysis and synthesis of nucleoside) reactions were studied individually. Thereby the role of various factors to achieve high yields of the desired nucleosides is discussed.

Chapter 4 contains the full paper (Paper III) about the successful immobilization of the thermostable NPs on the magnetic epoxide microspheres and the use of these biocatalysts for the small scale synthesis.

Finally, Chapter 5 (the summary and outlook of the Paper IV) demonstrates the immobilized NPs as biocatalysts for nucleosides synthesis are scalable (from 0.5 mL to 50 mL) and one of the important nucleoside precursors (6Cl-2F-purine riboside) was synthesized and isolated in high purity.

To be mentioned is that the "materials and method" part is written in Paper I~IV and throughout the thesis where is necessary.

Content

Abstract ..... **i**
Zusammenfassung ..... **ii**
Acknowledgements ..... **iv**
Original papers ..... **vi**
Abbreviations ..... **vii**
Preface ..... **x**
Content ..... **xii**

Chapter 1. Enzymatic synthesis of nucleosides (review) **1**
1.1. Enzymes in the synthesis of nucleosides *2*
1.1.1. NP and NDT (transglycosylation) 2
1.1.2. PPM and RK et al. (cascade biotransformation) 4
1.1.3. ADA, AMPDA, OX (modification) 7
1.2. Nucleoside phosphorylases as biocatalysts *7*
1.2.1. Enzyme source 7
1.2.2. Classification & structure 8
1.3. Modes of NP application *11*
1.3.1. Whole cells as natural encapsulated enzymes 11
1.3.2. Recombinant enzymes 12

Chapter 2. From gene sequences to the active NPs **14**
2.1. Enzyme source *14*
2.2. Recombinant expression *16*
2.2.1. Expression vectors 17
2.2.2. Expression optimization 19
2.2.3. Enzyme purification and production yield 25
2.3. Characterisation *27*
2.3.1. Temperature optimum 27

2.3.2. Stability 28
2.3.3. Kinetic properties and substrate specificities 31
2.3.4. pH optimum 36
2.3.5. Bioinformatics 38

Chapter 3. Transglycosylation processes 48

3.1. Type1: Equilibrium-controlled reactions 50
3.1.1. One-pot transglycosylation 50
3.1.2. Single step of the transglycosylation 53
3.2. Type2: Reactions which involve unstable intermediates 59
3.2.1. One-pot transglycosylation 59
3.2.2. Single step of the transglycosylation 60
3.3. Type3: Kinetically controlled reactions 62

Chapter 4. Immobilization of thermostable NPs (paper III) 65

4.1. Experimental 67
4.1.1. General 67
4.1.2. Procedures of enzyme immobilization 67
4.1.3. Enzyme activity assay 68
4.1.4. Immobilization yield (expressed activity) 69
4.1.5. Screening of immobilization conditions 69
4.1.6. Temperature profile of immobilized GtPNP 70
4.1.7. Thermostability of immobilized enzymes 70
4.1.8. Synthesis of 2,6-dihalogenated purine nucleosides 70
4.2. Results 71
4.2.1. Screening of immobilization conditions 71
4.2.2. Temperature profile of the immobilized GtPNP 76
4.2.3. Thermostability of immobilized enzymes 77

4.2.4. Synthesis of 2,6-dihalogenated purine nucleosides 78

4.3. Discussion 80

4.4. Conclusion 84

Chapter 5. Enzymatic synthesis (paper IV) 86

Chapter 6. Conclusion and outlook 89

6.1. Enzyme expression and characterisation 89

6.2. Transglycosylation processes 90

6.3. Immobilization 91

6.4. Enzymatic synthesis 91

6.5. Outlook 92

Chapter 7. Appendix 93

7.1. Multiple sequence alignment 93

7.1.1. Original alignment result of PyNP & TP from 3D-Coffee 93

7.1.2. Recoloured alignment of PyNP and TP 95

7.1.3. Recoloured alignment of PNP (I/II) and UP 96

7.2. 3D-structure of NPs by homology modelling 98

7.2.1. Structures of PyNPs 98

7.2.2. Structures of PNPs 99

7.2.3. Structures of UPs 100

7.2.4. Structure of MTAP 100

7.3. Supporting information (paper III) 101

References 103

Chapter 1. Enzymatic synthesis of nucleosides (review)

Nucleosides, composing of a nucleobase (A, T, C, G or U) and a pentofuranose (ribose or 2-deoxyribose), are the fundamental units of the genetic material (DNA and RNA) for all the living beings. Nucleosides and their derivates also take part in various biological processes. It has been believed that the synthesis of nucleosides and their analogues would be "the key to understanding the cellular functions and mechanisms involved in a variety of diseases." (Merino 2013) Indeed, nucleoside analogues have been extensively used as drugs for the treatment of cancers, viral / bacterial / fungal infections (Robak et al. 2009; De Clercq 2009; Böttcher and Sieber 2010; Kirk 2006); as precursors in the preparation of artificial oligonucleotides for therapeutic or diagnostic use (Watts and Corey 2012; Watts and Damha 2008); and as tools for investigation of mechanisms of enzyme function in molecular biology (Seela et al. 2009).

Nucleoside analogues have been synthesized traditionally by chemical methods through multistep processes requiring protection and deprotection steps for the labile groups, and tedious isolation in almost every step due to the poor *regio*- or *stereo*- selectivity of the reactions (Anderson et al. 2008; Cen and Sauve 2010; Tennilä et al. 2000). These drawbacks lead to a high price of the nucleosides and impeding their biological trials and studies, as well as limiting the wide therapeutic application (Mikhailopulo and Miroshnikov 2013).

In contrast, enzymatic methods are taken as an advantageous approach over the chemical methods, namely, with simple operation of reaction and product isolation, while result in high selectivity and efficiency. In additionally, compared with the fine organic synthesis, enzymatic methods to a greater extent meet to the demands of "green chemistry" (Mikhailopulo and Miroshnikov 2010).

In fact, enzymatic methods are "chemo-enzymatic methods" as the substrates of the enzymes have to be prepared by chemistry methods. In generally, substrates of the enzymes are simple compounds and commercial available. The products - nucleoside analogues - are often unavailable because they are difficult to synthesize by chemical methods.

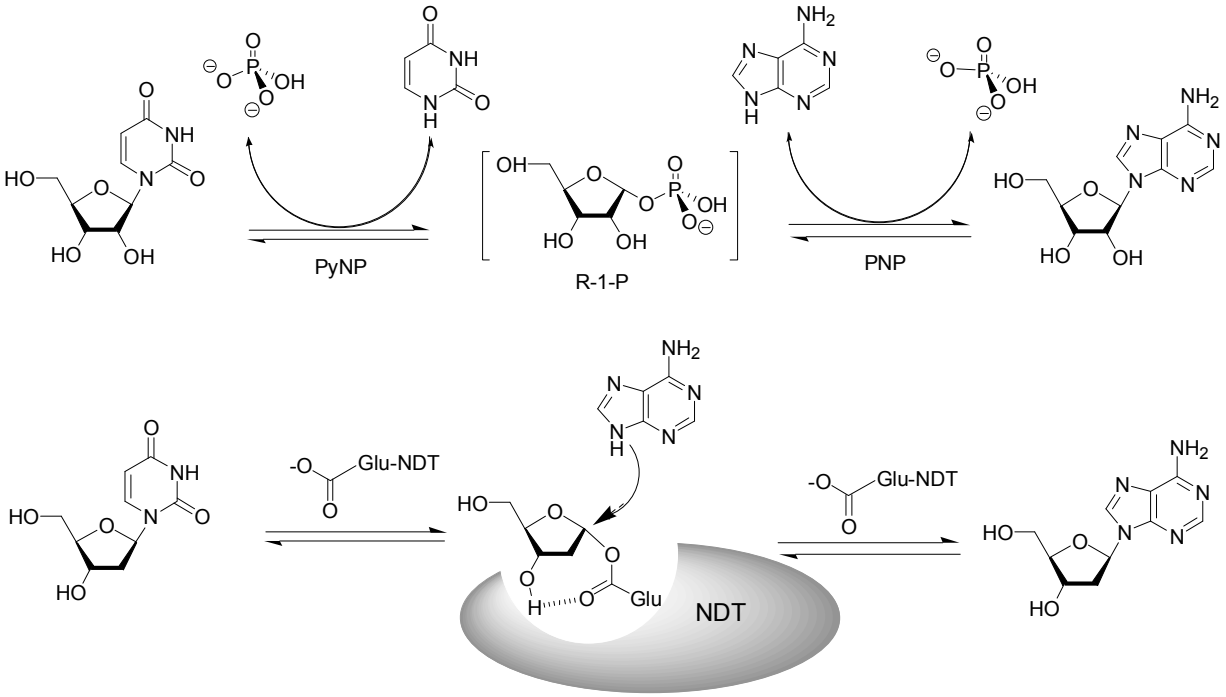
1.1. Enzymes in the synthesis of nucleosides

Different kinds of enzymes can be applied in the preparation of nucleoside analogues. Along with the synthesis strategies, the enzymes can be divided into three groups: (i) enzymes catalysing transglycosylation reactions, e.g. nucleoside phosphorylase (NP) and nucleoside deoxyribosyltransferase (NDT); (ii) enzymes used in the cascade transformation of pentose into the intermediate for nucleoside synthesis, e.g. ribokinase (RK) and phosphopentomutase (PPM); (iii) other enzymes used for the local modification of nucleosides, e.g. adenosine deaminase (ADA), adenylate deaminase (AMPDA), and xanthine oxidase (XO). The enzyme properties and their scopes of application will be specified as follows.

1.1.1. NP and NDT (transglycosylation)

Nucleoside phosphorylases (NPs) including pyrimidine NPs (PyNP, UP and TP) and purine NPs (PNP and MTAP), are the most used enzymes for the synthesis of modified nucleosides (Mikhailopulo and Miroshnikov 2011; Mikhailopulo 2007). These enzymes catalyse the reversible phosphorolysis of nucleosides in the presence of phosphate to form the corresponding α-D-ribofuranose-1-phosphate (R-1-P) and nucleobase. NPs display fairly broad substrate specificity, accept modified base and sugar moiety as substrates. The strategy of the transglycosylation reaction consists in the transfer of a sugar moiety from one nucleoside to another purine base, and then a new nucleoside is formed. The basic synthetic strategy is illustrated in Scheme 1.1. More details about NPs will be discussed later.

Differing from NPs, nucleoside deoxyribosyltransferases (NDT, EC 2.4.2.6) catalyse the direct transfer of the 2'-deoxyribosyl moiety from a 2'-deoxyribosyl nucleoside to a free nucleobase without formation of intermediate 2-deoxyribofuranose-1-phosphate (dR-1-P), and phosphate in not involved. Scheme 1.1 shows the transglycosylation reaction of NDTs. Specially, NDTs from *Lactobacilli* are well documented but their 3D structures haven't been expensively studied and only few entries are available from PDB.



Scheme 1.1. Transglycosylation reactions by NPs (above) and NDT (below).

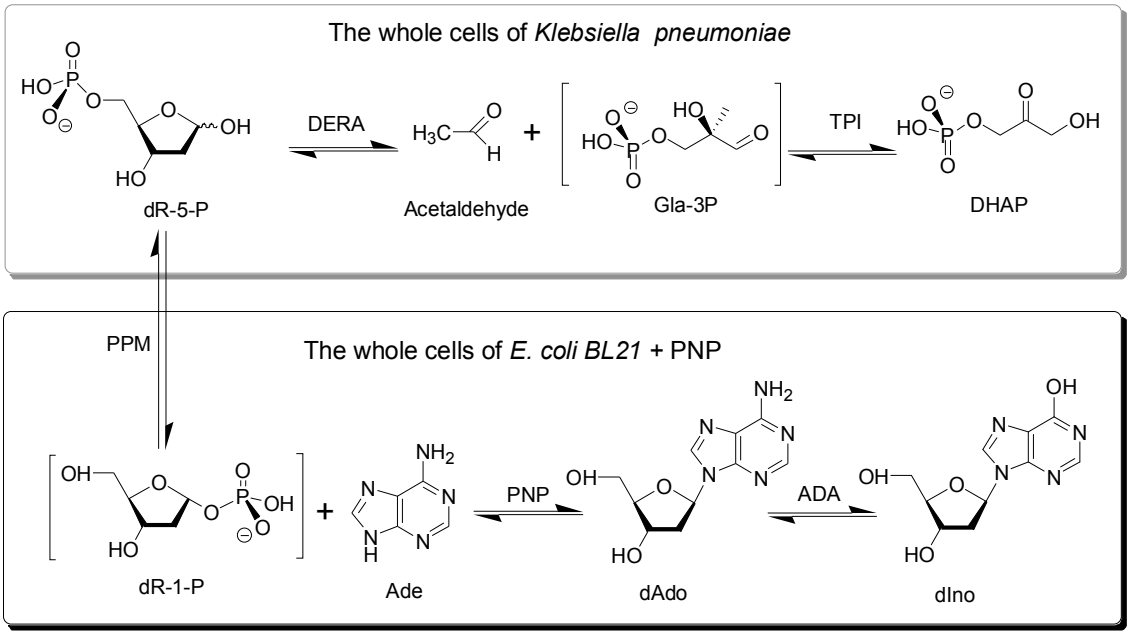
NDTs are classified into two types, NDTI and NDTII, based on their substrate specificity. NDTI is specific for purines; NDTII doesn't distinguish purines and pyrimidines (Kaminski 2002) but select cytosine as the best acceptor of 2'-dexoyribose moiety in all cases (Fresco-Taboada et al. 2013), while NPs have no or very low activity towards cytosine or cytidine. NDTs accept a broad range of modified bases, including different azole derivatives, size expanded purines, deaza- and halogenated purines and pyrimidines (Hutchinson 1990; Betbeder et al. 1989; Holguin and Cardinaud 1975; Fernandez-Lucas et al. 2010). NDTs have a strict specificity towards 2'-deoxyribonucleosides because of the steric hindrance that 3'-OH of the sugar moiety interacts with the catalytic residue Glu for the substrate orientation and catalysis (Fresco-Taboada et al. 2013) (see Scheme 1.1). However, some studies have revealed that the NDT from *Lactobacillus reuteri* exhibits unexpected new *arabinosyl-* and *2'-deoxy-2'-fluororibsyl*-transferase activities (Fernandez-Lucas et al. 2010); the NDT from *Lactobacillus helveticus* shows activity towards 2',3'-dideoxynucleosides (Carson and Wasson 1988); and the engineered NDT obtained an improved activity in regards to the synthesis of 2',3'-dideoxy and 2',3'-didehydro-2',3'-dideoxy purine nucleosides (Kaminski et al. 2008). Thus, in some degree, NDT may have as broad substrate specificity as NPs.

From the viewpoint of practical synthesis, both NDTs and NPs are potent biocatalysts and each of them has unique properties which can complement each other.

1.1.2. PPM and RK et al. (cascade biotransformation)

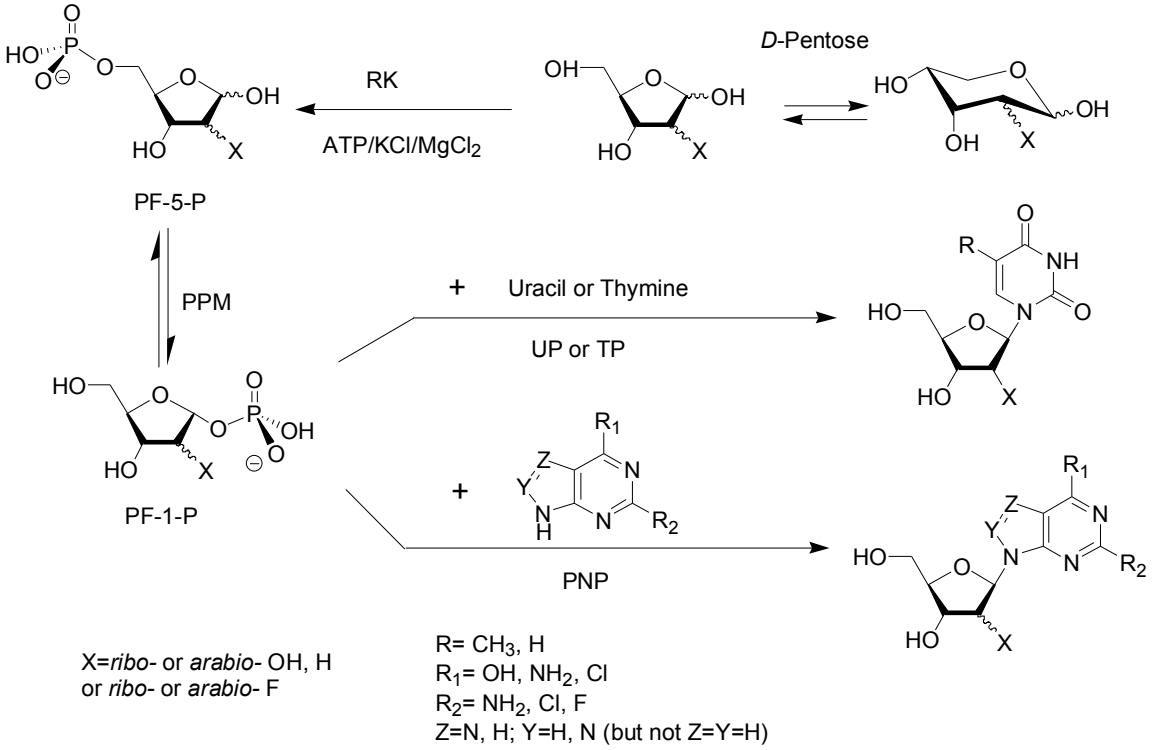
In *E. coli*, the *deo* operon consists of four structural genes: *deoC*, deoxyriboaldolase (DERA, EC 4.1.2.4); *deoA*, TP; *deoB,* phosphopentomutase (PPM, EC 2.7.5.6 or 5.4.2.7) and *deoD*, PNP. Interestingly, all of them could and have been used as biocatalysts for the synthesis of nucleosides in the cascade transformation from 2~3 carbon compounds to nucleoside (Ouwerkerk et al. 2002; Ishige et al. 2005). It is a good bionics example that we learn from the cell to use the simple starting materials (e.g. acetaldehyde, *D*-glyceraldehyde-3-phosphate and nucleobase) to synthesize nucleosides. The basic strategy is illustrated in Scheme 1.2.

Comparing with the cell metabolic transformations of pentoses and the *de novo* and salvage synthesis of nucleosides, the *in vitro* nucleoside synthesis is actually a reversed pathway (*retro*-pathway); therefore it is necessary to add great excess of acetaldehyde to direct the metabolic reaction in the reverse synthetic direction, thus a special strain *Klebsiella pneumonia* which tolerates high concentration of acetaldehyde was used. Moreover, this method is limited to the insufficient production of dR-1-P and much depending on the solubility of nucleobase. The off-pathway activities of the cells also prevent the efficiency of the desired product formation (e.g. in the cell, product dAdo was converted into dIno by adenosine deaminase, see Scheme 1.2). For a comprehensive review see (Mikhailopulo and Miroshnikov 2010).



Scheme 1.2. Retro-pathway for nucleoside synthesis by the whole cells.
Enzymes: *TPI= triose phosphate isomerise; DERA= deoxyriboaldolase; PPM= phosphopentomutase; PNP= purine nucleoside phosphorylase; ADA= adenosine deaminase.* ***Compounds:*** *DHAP= dihydroxyacetone monophosphate; Gla-3P= D-glyceraldehyde-3-phosphate; dR-5-P= 5-phosph-D-ribofuranose; dR-1-P= 2-deoxyl-a-D-ribose-1-phosphate; Ade= adenine; dAdo= 2'-deoxyadenosine; dIno= 2'-deoxyinosine.*

Recently, Mikhailopulo and co-workers proposed a novel nucleoside synthesis strategy of the cascade transformation (Mikhailopulo and Miroshnikov 2013; Chuvikovsky et al. 2006). It starts from D-pentose, which is phosphorolysed by ribokinase (RK, EC 2.7.1.15) into D-pentose-5-phosphate (PF-5-P), and the latter is converted to α-D-pentofuranose-1-phosphate (PF-1-P). The process is illustrated in Scheme 1.3. Noteworthy, the recombinant *E. coli* RK is able to phosphorolyse a set of *D*-pentofuranoses by which *D*-ribose, 2-deoxy-*D*-ribose, *D*-arabinose or 2-deoxy-2-fluoro-*D*-arabinose have been successfully transformed into the respective nucleosides in good yields. In this way, the anti-leukemic drugs Cladribine, Fludarabine, Clofarabine and Nelarabine (see Figure 2.9 for their structures), and a number of biologically important nucleosides (e.g. 8-azapurine and 8-aza-7-deazapurine 2'-deoxyriboncleosides) have been synthesized and the pure products have been obtained (Mikhailopulo and Miroshnikov 2013; Stepchenko et al. 2012). No doubt it is a simpler and more efficient method *versus* the retro-pathway in the cell.



Scheme 1.3. Cascade one-pot enzymatic transformation from pentose to nucleosides. Pure recombinant E. coli enzymes: RK=ribokinase; PPM= phosphopentomutase; PNP=purine nucleoside phosphorylase; UP or TP= uridine or thymidine phosphorylase.

It can be inferred that PF-1-P is the universal glycosylation agent for the enzymatic synthesis of pyrimidine and purine nucleoside. For this reason, various attempts have been made to synthesize PF-1-P by chemical way. However, the laborious and low-yielding preparation is the common bottleneck of the chemical approach. A relatively simple method proposed by D.L. MacDonald in 1960s has been taken as the most effective one. Recently Konstantinova and co-workers modified the MacDonald method and successfully combined it with the enzymatic method (using *E. coli* PNP or UP) for the preparation of purine and pyrimidine β-*D*-arabinofuranosides (Konstantinova et al. 2011).

It is worthy to point out that NPs play the key role in the abovementioned approaches. It was found that *E. coli* TP and UP show no activity towards some PF-1-Ps, e.g. 2-deoxy-2-fluoro-α-D-arabinofuranose-1-phosphate, thus the corresponding pyrimidine nucleosides cannot be synthesized by this method. Therefore, new NPs are demanded for the preparation of desired nucleosides.

1.1.3. ADA, AMPDA, OX (modification)

Other enzymes e.g. adenosine deaminase (ADA), adenylate deaminase (AMPDA), xanthine oxidase (XO) have been also used for the positional modification of nucleosides. They usually display broad substrate specificity and function on the special group of the nucleosides. For example, ADA catalyses the deamination of 6-aminopurines; AMPDA displaces 6-substitution (-NH_2, -OCH_3, -Cl) of purine by hydroxyl group; XO catalyses regioselective oxidations of numerous azaheterocycles (on C6 of purine). Since these enzymes are used for special purpose (e.g. deamination, acylation, deacylation halogenation) and are out of the scope of this thesis (see reviews (Li et al. 2010; Sinisterra et al. 2009)), no details will be further discussed here.

1.2. Nucleoside phosphorylases as biocatalysts

As it is shown either in the transglycosylation, or in the cascade transformation approaches, nucleoside phosphorylases (NPs) play the key role and are the most widely used biocatalysts for synthesis of modified nucleosides. To utilize NPs as biocatalysts, the enzyme source of the potential NPs for synthesis, their classification and their structures and catalytic mechanisms will be introduced.

1.2.1. Enzyme source

NPs are ubiquitous enzymes and have been found in bacteria, eukaryota, Archaea, and even in dsDNA viruses (Kang et al. 2010). Usually, an organism needs at least two kinds of NPs, i.e. PyNP (UP or TP) and PNP, which catalyse the phosphorolysis of pyrimidine and purine nucleosides, respectively.

A number of NPs or the original cells have been reported to be used as biocatalysts. A few of them are from thermophiles or with stable NPs above 50 °C, which include but are not limited to *Geobacillus stearothermophilus* (previously *Bacillus stearothermophilus*) (Hori et al. 1991a; Taran et al. 2009), *Enterobacter aerogenes* (Utagawa et al. 1985; Wei et al. 2008b), *Thermus thermophilus* (Almendros et al. 2012), and the hyperthermophilic Archaea *Aeropyrum pernix* (Zhu et al. 2012b), *Sulfolobus solfataricus* (Cacciapuoti et al. 1994; Cacciapuoti et al. 2005) and *Pyrococcus furiosus* (Cacciapuoti et al. 2003).

Thermozymes have special advantages and recently attract much attention of researchers. From the viewpoint of practical synthesis, thermozymes offer the possibility to perform reactions at high temperature resulting in decrease of the viscosity of the medium, increase diffusion coefficients of substrates and lead to higher overall reaction rates. The product yield may be improved due to the higher solubility of substrates at the high temperature (Bruins et al. 2001). Moreover, thermozymes are often resistant to pressure and organic solvents (Hei and Clark 1994). They can be expressed in *E. coli* at high levels and easily purified by heat treatment (Almendros et al. 2012). Therefore, thermostable NP is the focus for this study.

1.2.2. Classification & structure

Based on the enzyme structures, NPs are classified into two big families: NPI and NPII (Pugmire and Ealick 2002). The NPI family contains enzymes which share a common single-domain subunit with a trimeric or a hexameric quaternary structure, which includes purine nucleoside phosphorylase (PNP, EC 2.4.2.1), 5′-methylthioadenosine phosphorylase (MTAP, EC 2.4.2.28) and uridine phosphorylase (UP, EC 2.4.2.3). Enzymes from the NPII family share a common two-domain subunit fold and a dimeric structure; they include pyrimidine nucleoside phosphorylase (PyNP, EC 2.4.2.2) and thymidine phosphorylase (TP, EC 2.4.2.4).

Despite of the wide range of substrate specificities and low sequence identity among the enzymes of the NPI family, they share a conserved structure which distinguishes them from the NPII family. The subunit fold of NPI enzymes contains a distorted β-barrel surrounded by several α-helices (Figure 1.1 A) while the subunit fold of NPII enzymes consists a large α/β-domain and a smaller α-helical domain (Figure 1.1 B).

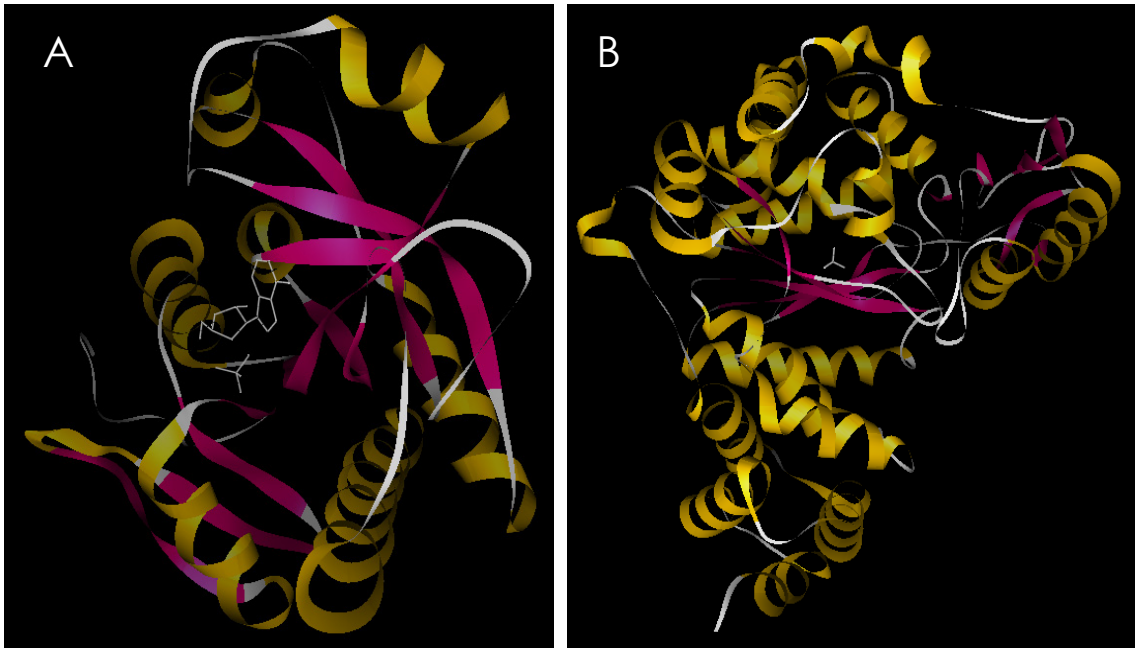


Figure 1.1. Structure of NPI and NPII family enzymes. Ribbon drawing of the subunit of (A) EcPNP (PDB: 1P7W, with adenosine and phosphate) and (B) EcTP (PDB: 4EAF, with SO_4^{2-}).

From the substrate point of view, PNP and MTAP of the NPI family can be divided into PNPI and PNPII subfamilies and the former has a broader substrate specificity and is specific for both 6-oxopurine and 6-aminopurine nucleosides; the latter is only specific for 6-oxopurine nucleosides. Empirically, PNPI has a hexameric structure or a high molecular mass (110~160 kDa) which mainly exists in bacteria; PNPII has a trimetric structure or a low molecular mass (80~100 kDa) and mainly exists in mammals and some bacteria (Bzowska et al. 2000). The typical representatives are *E. coli* PNPI and PNPII from *deoD* and *xapA* gene products, respectively. MTAP was categorized as PNPII, probably due to the fact that the earlier reported MTAPs were isolated from mammals. Recently, a number of MTAPs from hyperthermophilic Archaea have been characterised, which are from *A. pernix* (paper II), *S. solfataricus* (Cacciapuoti et al. 1994; Cacciapuoti et al. 2005) and *P. furiosus* (Cacciapuoti et al. 2003; Cacciapuoti et al. 2007). Like many bacteria, Archaea express two kinds of PNPs (MTAPI and MTAPII) with different substrate specificities, but their oligomeric states are unclear. For utilizing these enzymes as biocatalysts, it is more practical to classify them according to their substrate specificities, namely, (i) specific for 6-oxo- and 6-aminopurines (most PNPI members); (ii) specific for 6-oxopurines (most PNPII members, TtPNPI as

an exception); (iii) hightly specific for 6-aminopurines (e.g. Ado-PNP from *Bacillus cereus* and *Bacillus subtilis*, SsMTAPII, TtPNPII).

UP (hexamer) is very special because it has a similar structure as NPI family members but shares the same substrate as PyNP from NPII family, i.e. UP and PyNP catalyse pyrimidine nucleosides (uridine and thymidine) and do not distinguish ribose and 2'-deoxyribose moieties, while TP is specific for 2'-deoxynucleosides and does not distinguish thymine and uracil.

PyNP and TP from the NPII family share a high degree of sequence identity (above 30 %) and they require significant domain movements for substrates binding, which leads to the conformation change of domains from open to close state (Pugmire and Ealick 1998). This is an important property to be considered in immobilization, since rigid supports may hinder the domain movements and result in an activity lost (Serra et al. 2013b).

Interestingly, iterative BLAST searches with various members of NPI family reveal great similarities between the subunit of the hexameric NPI members and both MTA/SAH and AMP nucleosidases (Pugmire and Ealick 2002). They all cleave the glycosidic bond, but nucleosidases use water as the nucleophile while phosphorylases (NP) use phosphate as the nucleophile. On the other hand, it suggests a possible structural relationship between NPII (PyNP) and anthranilate phosphoribosyltransferase (AnPRT, EC 2.4.2.18) due to the similarities of their catalysed reactions, the properties of the reactants and products, and their 2D and 3D structures (Pugmire and Ealick 2002; Mushegian and Koonin 1994). These data can be used for analysis of the potential substrate specificity of NPs.

Although big differences exist in the 2D and 3D structures of NPs, the basic catalytic mechanism of their phosphorolysis is presumed to be similar, which undergoes (i) nucleoside binding in a high energy conformation in order to produce steric strain for the glycosidic bond breaking; (ii) an oxocarbenium ion formation through electron movement from O4' of the pentose moiety to the purine ring; (iii) a nucleophilic attack at C1' of the pentose by a phosphate ion. Finally, the glycoside bond is broken and phosphate is connected to the C1'

position in α-configuration, i. e. an α-D-Ribose-1-phosphate is formed and a nucleobase is released (from the sugar moiety).

1.3. Modes of NP application

There are mainly two modes of using NPs in synthesis applications: (i) the whole cells as natural encapsulated enzymes; and (ii) recombinant enzymes. As the whole cells or the enzymes can be used as extracts or in an immobilized form, there are in total four modes of how NPs are applied. Examples of each mode will be given to analyse their advantages and disadvantages.

1.3.1. Whole cells as natural encapsulated enzymes

E. coli cells after treatment or modification at gene level can be applied as biocatalysts for the synthesis. For example, Barai and co-workers used glutaraldehyde treated *E. coli* cells BMT 4D/1A as biocatalysts for the synthesis of Cladribine (Barai et al. 2002), with 2Cl-Ade as an acceptor and dGuo as a donor of 2-deoxyribofuranose (1:3 mol ratio) of the transglycosylation reaction. In this heterogeneous reaction mixture (10 mM *K*-phosphate buffer, pH 7.5; 65°C), 2Cl-Adewas converted in two steps (4 h reaction followed by the separation of substrates and biocatalyst which were both incubated in fresh buffer for additional 3 h) to obtain finally, after silica gel column chromatography, pure Cladribine in a yield of 81% based on 2Cl-Ado taken into reaction.

Another example is about screening of different strains from *Thermus thermophilus* (Almendros et al. 2009). The pellet from the harvested cells was tested in order to detect the transglycosylation activity using (d)Urd or Thd as the sugar donor and adenine or hypoxanthine as bases (1:1 mole ratio) in 30 mM phosphate or Tris/HCl buffer at 65 °C for 1 h. The author supposed *T.thermophilus* may contain NDT as a higher activity was observed with 2'-deoxynucleoside as sugar donor and the reaction was performed in absence of phosphate.

Recently, Nobile and co-workers (Nóbile et al. 2012) selected *Citrobacter koseri* (CECT 856) from a collection of microorganisms as the best biocatalyst for the preparation of purine arabinosides. The selected whole cells were immobilized in agarose (catalyst load: 1.4×10^{11} cells g^{-1} to prepare arabinonucleosides) and

used as biocatalysts for the synthesis of three drugs at 60 °C from 0.05 mL to 150 mL enlarged scale with 3 mM 2-F-Ade and 6 mM AraU for fludarabine, 10 mM adenine and 20 mM AraU for vidarabine, or 10 mM 2,6-diamiopurine (DAP) and 30 mM AraU (in 30 mM potassium phosphate buffer, pH 7) with 58 % yield in 14 h, 71 % yield in 26 h, and 77 % yield in 24 h, respectively. This work proved that the biotransformation is scalable since no significant changes were observed in the profile of the reactions from 0.05 mL to 150 mL.

1.3.2. Recombinant enzymes

With molecule cloning techniques, many enzymes from the microorganisms which are hard to grow under the lab conditions (e.g. hypothermophilies) have been recombinantly expressed in *E. coli* (Zhu et al. 2012a; Almendros et al. 2012; Cacciapuoti et al. 2007). Recombinant enzymes are often used without immobilisation as free enzymes for the enzyme characterisation and down-stream optimization; however if they are applied in the synthesis of nucleosides they need to be immobilized to reduce biocatalyst cost and to increase enzyme recovery.

Taran et al (Taran et al. 2009) immobilized recombinant PyNP and PNP from *G. stearothermophilus* on aminopropylated macroporous glass AP-CPG-170 and used them as a biocatalyst in the transglycosylation reaction (5 mM *K*-phosphate buffer, pH 7.5; 70-75 °C; 2-8 h) for the synthesis of Cladribine (thymidine & 2Cl-Ade as substrates with a mol ratio of 3:1) and Fladarabine (AraU and adenine as substrates with a mol ratio of 3:1) in the reactions containing 10-20 % DMSO (for increasing solubility of substrates). Analysis of the reaction mixtures by HPLC showed the conversion of the respective bases to Cladribine and Fladarabine with a yield of 86 and 88 %, respectively. The yields are very promising, however, the low solubility of the substrates is the drawback and the use of DMSO causes difficulties in the separation and purification of the final product.

In summary, for the synthesis of nucleosides, especially for 2'-deoxynucleosides and arabinonucleosides, the enzymatic method has obvious advantages over the chemical method in many ways. A rational combination of chemical approach (e.g. MacDonald method for the preparation of PF-1-P) and

enzymatic methods (by immobilized NP or NDT) are the new trend for the application of biotechnology in fine chemistry.

Chapter 2. From gene sequences to the active NPs

In this chapter, six target enzymes were selected for the investigation based on two criteria: 1. the candidate enzymes are intrinsic stable; 2. the enzymes should have as broad substrates specificity as possible. Thanks to the nature's abundant gene resources, the enzymes from the microorganisms which are able to survive and even thrive at high temperatures (thermophiles) would fulfil the above requirements. Therefore, taking advantage of the bioinformatics and cloning techniques, strategies for the production of six promising enzymes (GtPyNP, TtPyNP, DgPNP, GtPNP, ApMTAP and ApUP) were developed, and the first five enzymes were successfully expressed, purified and characterised.

The main work and the results[1] on the subject of enzymes expression and characterisation were published in paper I (GtPyNP and TtPyNP) (Szeker et al. 2012) and paper II (DgPNP, GtPNP and ApMTAP) (Zhou et al. 2013). Materials and methods are consistent with the papers mentioned above if not otherwise stated. A summary of the enzyme source, an overview of the related gene cloning, enzyme expression and their physicochemical properties as well as the characterisation results will be provided; new insights on bioinformatics will be also given hereafter.

2.1. Enzyme source

The thermophilic bacteria *Deinococcus geothermalis, Geobacillus thermoglucosidans* (formerly *Bacillus thermoglucosidasius* (Suzuki et al. 1983) and *G. thermoglucosidasius* (Nazina et al. 2001)), *Thermus thermophilus,* and the hyperthermophilic archaeon *Aeropyrum pernix* were the microbial sources for the target gene isolation. The desired enzymes and the information of the species are summarized in Table 2.1.

[1] Most of the cloning work (except TtPyNP) was done by Kathleen Szeker (previously PhD student, Neubauer group, TU Berlin). The expression, optimization and characterisation of the enzymes was done in close cooperation with K. Szeker.

Table 2.1. Thermophilic organisms as the source for the target enzymes in this study.

Organism	*Deinococcus geothermalis*	*Geobacillus thermoglucosidans*	*Thermus thermophilus* HB 27	*Aeropyrum pernix*K1
Reference	(Ferreira et al. 1997)	(Suzuki et al. 1983)	(Oshima and Imahori 1974)	(Sako et al. 1996)
Temperature optimum [°C]	30~55 (45~50)	42~69 (61~63)	47~85 (65~72)	68~102 (90~95)
pH (opt.)	4.5~8.5 (6.5)	6.5~8.5 (6.5)	5.1~9.6 (7.0)	5.0~9.0 (7.0)
Characteristics	Strictly aerobic, extreme radiation resistance, slightly acidophilic, able to grow in minimal medium without yeast extract	Strictly aerobic, obligate thermophilic spore-former, neutrophilic, amylolytic	Strictly aerobic, obligate thermophile, nonsporulating, lysozyme resistant below 60 °C, proteolytic	Strictly aerobic, hyperthermo-philic archaeon, highly developed thiosulfate-oxidizing heterotrophs
Gram staining	positive	positive	negative	negative
C+G cont. [mol %]	65.9	43.9	69.4	67
Isolation place	Hot springs, Italy	Soils, Japan	Hot springs, Japan	Coastal solfataric vent, Japan
Genome sequence	CP000856 (2007)	CP002835 (2011); (Zhao et al. 2012)	NC_005835 (Henne et al. 2004)	NC_000854 (Yamazaki et al. 2006)
Target enzymes	DgPNP	GtPNP, GtPyNP	TtPyNP	ApMTAP, ApUP

The reason of isolating enzymes from different thermophilic organisms is to have more opportunities to get the potential biocatalysts activity towards different substrates, on the one hand, and the different microorganisms having different optimal growth temperatures (Table 2.1), which might provide an enzyme with a lower temperature optimum especially suitable for the process involving thermolabile reactants, on the other hand.

As shown in Table 2.1, four thermophilic organisms demonstrate diverse characteristics, since they are phylogenetically far from each other. *D. geothermalis* and *T. thermophilus* are from the same phylum *Deinococcus-Thermus* as well as the same class *Deinococci*, but the former is in the order *Deinococcales* and the latter is in the order *Thermales*. From a different phylum

Firmicutes, *G. thermoglucosidans* is in the same genus as *G. stearothermophilus*, and the latter is the enzyme source of the thermostable NPs (GsPNP and GsPyNP) that have been intensively studied and proved to be promising biocatalysts (Rivero et al. 2012; Trelles et al. 2005; Taran et al. 2009; Hamamoto et al. 1996; Hori et al. 1991a; Hori et al. 1991b). Hence, the enzymes derived from *G. thermoglucosidans* (GtPNP and GtPyNP) are expected to have the desired properties. *A. pernix* is a distinct species from the third domain of life - Archaea (Pace 2006). Most of their metabolic enzymes are common to bacteria, while their proteins involved in gene expression are common to Eukarya, thus it was suggested that they might provide NPs with unique activities. Besides, a few reported NPs from Archaea are all hyperthermophilic enzymes, which are SsMTAP I/II derived from *Sulfolobus solfataricus* (Cacciapuoti et al. 2005; Cacciapuoti et al. 2001; Appleby et al. 2001; Cacciapuoti et al. 1994), PfMTAP and PfPNP from *Pyrococcus furiosus* (Cacciapuoti et al. 2011; Cacciapuoti et al. 2007; Cacciapuoti et al. 2004; Cacciapuoti et al. 2003), as well as ApMTAP and ApUP from *A. pernix* (Zhu et al. 2012a; Zhu et al. 2012b; Montilla Arevalo et al. 2011). These studies were mainly focussed on the basic research of the structure or the catalytic mechanisms of these enzymes as a basis for their later application. In this study hyperthermozymes ApMTAP and ApUP are selected although they have been considered for synthetic applications, their catalytic characteristics are not available (Zhu et al. 2012b), which prompt us to investigate them further.

2.2. Recombinant expression

Thermozymes become applicable biocatalysts only when they can be produced in their active form in an adequate amount. However, most thermophiles are difficult to grow under laboratory conditions. Using recombinant DNA technology, the target genes can be cloned into suitable hosts, e.g. in *E. coli*, which is the most convenient and extensively utilized organism for the expression of heterologous proteins with many related successful examples. Therefore, *E. coli* was chosen as the microbial host for the expression of the recombinant enzymes in this study. The detailed experiments and the results have been published in the thesis of K. Szeker (Szeker 2012) as well as in the papers I & II. The confronted challenges and the lessons learned through them will be discussed here.

2.2.1. Expression vectors

Initially, the DgPNP gene was cloned into the expression vector pCTUT7 (Siurkus et al. 2010), and later in a derivative of this plasmid, called pCTUT7A (Szeker 2012), by substituting the chloramphenicol resistance cassette with the plasmid stabilizing *parB* locus (Gerdes 1988) and an ampicillin resistance cassette (Figure 2.1). Both vectors contain an IPTG inducible *lac* promoter, a pBR322 origin of replication, and an upstream (in N-terminal) hexahistidine tag with additional amino acids of the *attB1* recombination site and the TEV protease cleavage site (Figure 2.2).

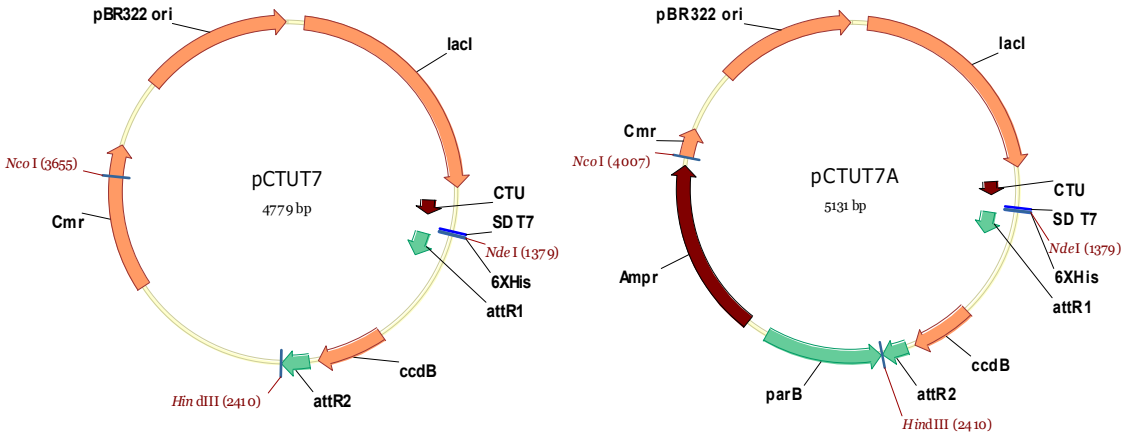


Figure 2.1. Vector maps of pCTUT7 and pCTUT7A.

DgPNP was only expressed in insoluble form with the pCTUT7 vector. To prevent DgPNP aggregation, a set of plasmids encoding chaperones (Takara plasmid pGro7, encoding groES-groEL, and pG-KJE8, encoding groES-groEL together with dnaK-dnaJ-grpE) was co-transformed into *E. coli* BL21 pCTUT7A_DgPNP. The results showed that the coexpression of chaperones improved the soluble DgPNP expression, but the protein couldn't withstand a 50 °C heat-treatment, which indicated that DgPNP was not expressed in its native conformation.

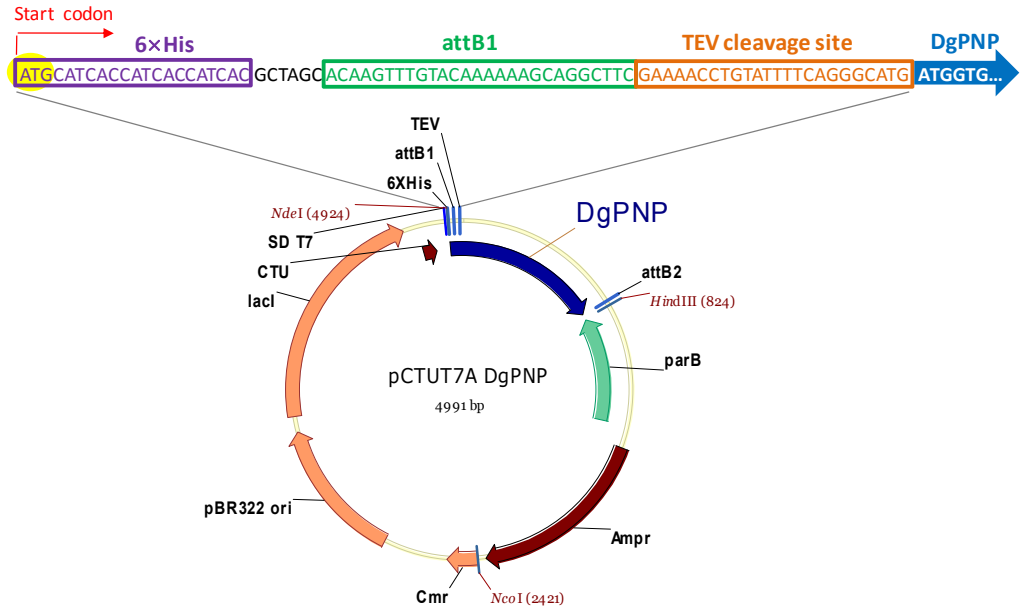


Figure 2.2. Map of the expression vector pCTUT7A_DgPNP.

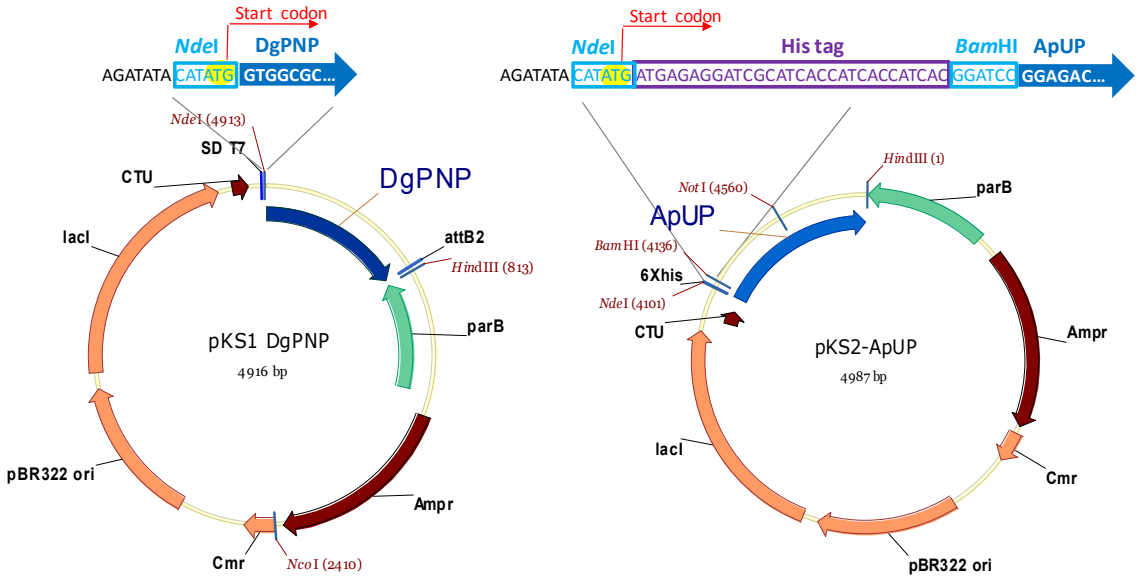


Figure 2.3. Map of the expression vectors of pKS1_DgPNP and pKS2_ApUP.

The extra amino acids in the "long" his-tag (Figure 2.2) were supposed to be the reason of wrong folding of the expressed protein. Thus two new expression vectors – one without his-tag (pKS1) and another one with a "short" his-tag (pKS2) - were constructed by cloning the DgPNP gene and the "short his-tag + *Bam*HI + ApUP" sequence into the pCTUT7A expression vector, respectively (Figure 2.3).

Thereafter, the other NP genes could be conveniently inserted into the vectors pKS1 and pKS2 to test the effect of this his-tag on protein expression.

Other expression vectors: pET21a (with C-terminal his-tag and without), pQE40 (with N-terminal his-tag), and pKS2 with co-transformed pMJS9 or pFH 255 for disulfide bond containing protein expression (Nguyen et al. 2011; Hatahet et al. 2010) were also tested. Since no obvious advantage was observed with them, the final expression vector was pKS2 for all the enzymes expressed in this study (Figure 2.4).

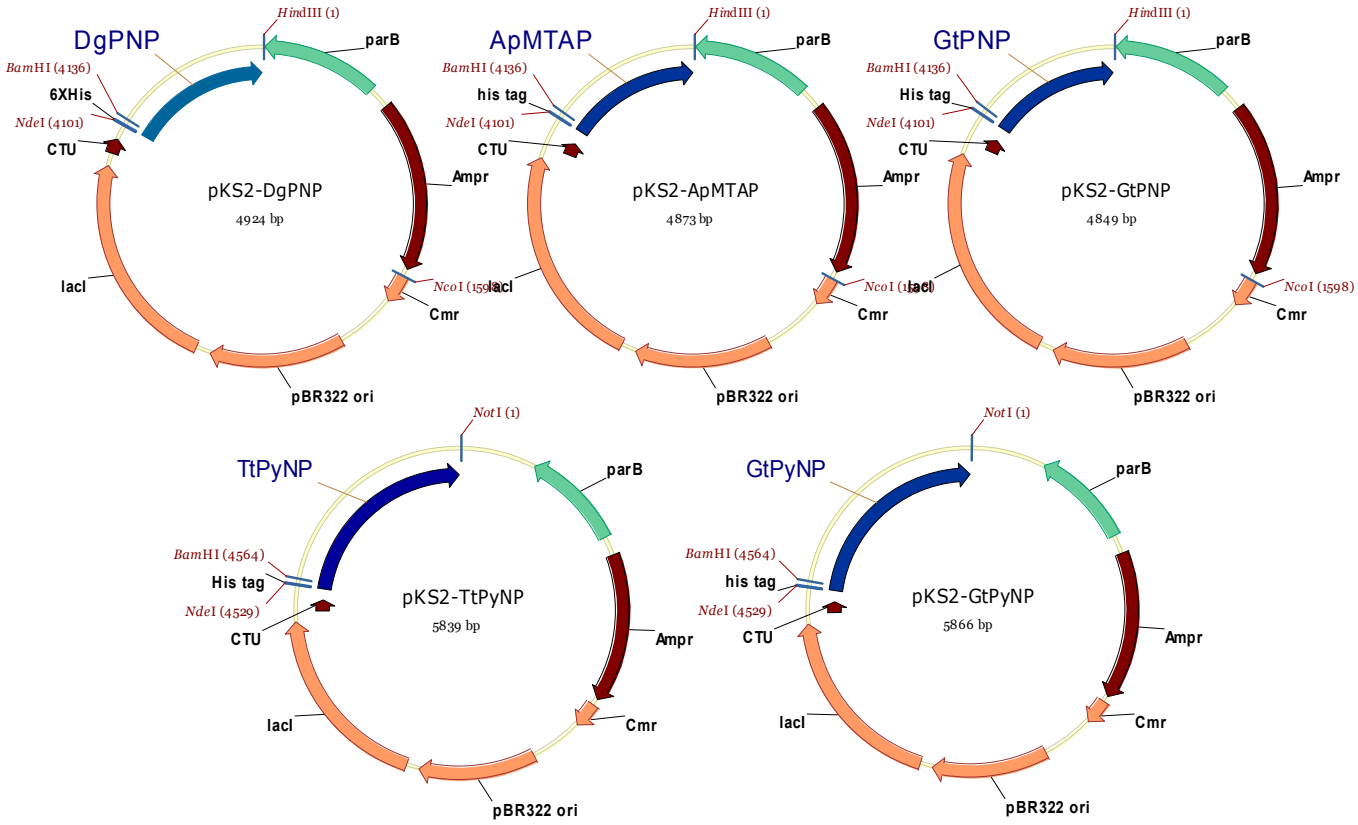


Figure 2.4. Final expression vectors of the target enzymes.

2.2.2. Expression optimization

The theoretically calculation (by Vector NTI software, Invitrogen) of the enzymes' physicochemical properties from the protein sequence are listed in Table 2.2 (without his-tag) and Table 2.3 (with his-tag).

Table 2.2. Theoretical analysis of the enzymes without his-tag.

Analysis	DgPNP	ApMTAP	GtPNP	GtPyNP	TtPyNP	ApUP
Length (aa)	261	244	235	432	424	283
Mw [kDa]	28.45	25.47	26.00	46.03	45.41	30.32
[mmol g^{-1}]	35.150	39.267	38.465	21.724	22.022	32.983
ε [$M^{-1}cm^{-1}$]	18730	23710	24420	19890	26150	38480
1 A_{280} corr. to [mg ml^{-1}]	1.52	1.07	1.06	2.31	1.74	0.79
A_{280} of 1 % [AU]	6.6	9.3	9.4	4.3	5.8	12.7
Isoelectric Point	5.4	6.0	5.3	5.2	6.1	6.4
Charge at pH 7	-6.7	-2.0	-5.9	-9.8	-4.4	1.0

Table 2.3.Theoretical analysis of the expressed his-tagged enzymes.

Analysis	*h*_DgPNP	*h*_ApMTAP	*h*_GtPNP	*h*_GtPyNP	*h*_TtPyNP	*h*_ApUP
Length (aa)	274	257	249	445	436	295
Mw [kDa]	29.96	26.98	27.62	47.56	46.81	31.74
[mmol g^{-1}]	33.379	37.070	36.205	21.025	21.364	31.511
ε [$M^{-1}cm^{-1}$]	18730	23710	24420	19890	26150	38480
1 A_{280} corr. to [mg ml^{-1}]	1.60	1.14	1.13	2.39	1.79	0.82
A_{280} of 1 % [AU]	6.3	8.8	8.8	4.2	5.6	12.1
Isoelectric Point	6.0	6.8	6.1	5.7	6.5	7.3
Charge at pH 7	-5.2	-0.4	-4.3	-8.3	-2.8	0.6

The enzyme expression conditions (including expression vector, expression strain, cultivation medium and temperature) and the results as well as the final expression conditions and results used for purification and characterisation are summarized in Table 2.4 – Table 2.9.

DgPNP

Table 2.4. Expression of DgPNP.

Vector_strain	Tag	Note	Results
BL21 pCTUT7	N-his-tag (long)	TB medium, 20~37 °C	insoluble
BL21 pCTUT7A	N-his-tag (long)	co-expression of chaperones, TB medium, 30 °C	soluble but not folded rightly (degraded > 50 °C)
BL21 pKS1	None	optimized 5'-mRNA (cultivation at 30~42 °C), TB medium	soluble and active; expression level increased at higher cultivation temperature (Szeker et al. 2011)
BL21 pKS2	N-his-tag (short)	TB medium, 37 °C	mostly insoluble; OD_{600}=16 (2.5 h)
		EnPresso® B, 37 °C	soluble and active; OD_{600}= 25 (24 h)

Final expression conditions: *E. coli* BL21 pKS2_DgPNP, EnPresso B medium, 37 °C, cells were harvested at 24 h after induction (20 μM IPTG). Final OD_{600}= 25.

ApMTAP

Table 2.5. Expression of ApMTAP.

Vector_strain	Tag	Note	Results
BL21 pKS1_ApMTAP BL21(Rosetta) pKS1_ApMTAP	None	optimized 5'-mRNA (cultivation at 30~42 °C), TB medium, 0~100 μM IPTG	soluble and thermostable; expression level increased at high cultivation temp.; Rosetta showed no advantage *vs* BL21; higher cell density was obtained in BL21
BL21 pKS2_ApMTAP	N-his-tag (short)	TB medium, 37 °C	mostly soluble; thermostable at 90 °C; OD_{600}=10 (2.5 h)
		EnPresso B, 37 °C	mostly soluble; active; OD_{600}=20 (22 h)

Final expression conditions: pKS2_BL21_ApMTAP, EnPresso medium, 37 °C, cells were harvested at 22 h after induction (20 μM IPTG). Final OD_{600}= 20.

GtPNP

Table 2.6. Expression of GtPNP.

Vector_strain	Tag	Note	Results
Rosetta2(DE3) pET21a_GtPNP	None	TB medium, 37 °C, 100 µM IPTG	high amount soluble fraction; thermostable at 65 °C
	C-his-tag		insoluble
BL21 pKS2_GtPNP	N-his-tag (short)		soluble; active; OD_{600}=16 (3.5 h)

Final expression conditions: *E. coli* BL21pKS2_GtPNP, TB medium, 37 °C, cells were harvested at 3.5 h after induction (100 µM IPTG). Final OD_{600}= 16.

GtPyNP

Table 2.7. Expression of GtPyNP.

Vector_strain	Tag	Note	Results
Rosetta2 (DE3)or Rosetta-gami(DE3) pET21a_GtPyNP	None	TB medium, 37 °C, 100 µM IPTG	high amount soluble fraction; not thermostable >45°C with BugBuster®
BL21 pKS2_GtPyNP	N-his-tag (short)		high amount soluble fraction; not thermostable>45°C with BugBuster, but thermostable at 60°C with sonication buffer (50 mM KP, pH 8.0)

Final expression conditions: *E. coli* BL21 pKS2_GtPyNP, TB medium, 37 °C, cells were harvested at 3.5 h after induction (100 µM IPTG). Final OD_{600}= 11.

TtPyNP

Table 2.8. Expression of TtPyNP.

Vector_strain	Tag	Note	Results
Rosetta2 (DE3) pET21a_TtPyNP	C-his-tag	30/37°C, LB/TB medium, 100 µM IPTG	soluble and thermostable at 80 °C
RB791 pQE40_KK8_TtPyNP	N-his-tag	30/37°C, TB/EnPresso B medium, 100 µM IPTG	soluble, thermostable and active; very low expression level; no advantage of using Rosetta
Rosetta2(DE3) or BL21pKS2_TtPyNP	N-his-tag (short)	TB/EnPressoB medium, 37 °C, 0.1-10 mM IPTG	

Final expression conditions: *E. coli*BL21 pKS2_TtPyNP (or Rosetta2(DE3) pET21a_TtPyNP), EnPresso B medium, 37 °C, cells were harvested at 24 h after induction (1 mM IPTG). Final OD_{600}= 24.

ApUP

Table 2.9. Expression of ApUP.

Vector_strain	Tag	Note	Results
BL21 pKS1_ApUP	None	37°C, TB medium,0~100 µM IPTG	small soluble fraction
Rosetta-gami pKS1_ApUP			big amount of insoluble fraction
Origami pKS1_ApUP			Insoluble fraction
BL21pKS1_ApUP		ApUP1 (remove the potential internal ribosomal binding site)	No visible improvement
BL21pKS2_ApUP	N-his-tag (short)	ApUP1(see above)	
		ApUPsh (84 bp shorter)	insoluble
BL21/Rosetta2/OrigamipKS2_ApUP		TB/EnPresso Bmedium, 30/37 °C, 0-100 µMIPTG	Rosetta strongly expressed but only in insoluble form; Origami and Rosetta-gami weakly expressed in soluble form
BL21 pKS2_ApUP +pMJS9(or pFH255)		Arabinose 0.5 % + 0.1 mM IPTG	Erv1p (pMJS9 or pFH 255)showed no improvement

ApUP was not successfully expressed in this study. The best results obtained from pKS1_BL21 with 20 µM IPTG in TB medium at 37 °C. However, ApUP was expressed by others with different tags (Table 2.10).

Table 2.10. Expression of ApUP from the literature.

Vector_strain	Tag	Note	Results
D-TOPO®_BL21 pET102	N-his-thioredoxin & C-his-tag	ApUPsh (84 bp shorter), 37 °C, 0.1 g L^{-1} IPTG, 8 h	thermostable at 80 °C for 10 h; resistant to 5 % DMSO (Montilla Arevalo et al. 2011)[a]
BL21(DE3) pET28a/30a	C-his-tag/None	37 °C, LB medium, 1 mM IPTG	only no-his-tag ApUP was active and soluble (hexamer), OD_{580}=1.2 (4 h), thermostable at 90 °C(Zhu et al. 2012b)[b]

[a] OD value is not available. A shorter ApUP gene sequence was applied, which was supposed that the sequence started after the internal ribosomal binding site.
[b] The same ApUP gene sequence as we used in this study.

On the whole, five of the six nucleoside phosphorylases from thermophiles or hypothermophiles were successfully expressed in *E. coli*, although in different yields. The main factors which need to be considered for the expression of thermozymes in *E. coli* could be summed up as follows.

- His-tag: with or without, position (N- or C-terminal) and its length (e.g. whether including TEV cleavage site or other extra amino acids) may be harmful to the protein folding, but it may be helpful for reducing the 5' mRNA structure stability to facilitate translation initiation (e.g. pKS2 functions in this way).
- 5' mRNA stability: A stable 5' mRNA structure of the thermozymes' gene can be easily formed at low cultivation temperature, which might impede the protein translation/expression. Two approaches were valid for alleviating this problem – increasing cultivation temperature and making silent mutations in the 5' mRNA region.
- Native enzyme or not: native enzyme should be expressed in soluble form; at the same time it should be resistant to the heat-treatment at least at the temperature of the origin cell's optimal growth.
- Disulfide bonds: The online tool CYSPRED predicted that two enzymes (ApMTAP and ApUP) from *A. pernix* contain one and two disulfide bonds, respectively (Szeker 2012). Origami and Rosetta-gami, two *E. coli* strains engineered to enhance disulfide bond formation in the cytoplasm, were

employed in this study. However, none of them showed significant improvement for the soluble protein expression compared with BL21. In addition, Origami and Rosetta-gami showed much slower growth rate than BL21, and they are not suitable for fermentation application.

- Rare codons: It was shown that the amount of rare codons in a gene sequence was related with the phylogenetic distance of the thermophile from *E. coli* (Szeker 2012). Indeed, the codon adaption index (CAI, a parameter used to predict gene expression levels, value "1" is optimal (Sharp and Li 1987)) decreased from 0.74 (DgPNP) to 0.5 (ApMTAP), and the expression level (in BL21) did become lower. Rosetta, an *E. coli* derivative designed to supply rare tRNAs for the expression of heterologous genes encoding many rare codons, was applied in all cases except for the expression of DgPNP. The results did not show any positive effects of the Rosetta strain. Thus BL21 was chosen as the expression strain for the final expression.
- Medium: The soluble enzymes were substantially increased when the controlled substrate delivery technology (EnPresso B medium) (Ukkonen et al. 2011; Krause et al. 2010) was applied especially in the case of DgPNP and ApMTAP. Another advantage of EnPresso B is that high cell-densities can be reached, and thus high volumetric yields of the enzymes were obtained.

2.2.3. Enzyme purification and production yield

Due to the his-tag and the thermostable property, the studied enzymes were purified by the combination of heat-treatment and immobilized metal affinity chromatography. The purified enzymes showed good purity (Figure 2.5). It is worthy to note that a second band of ApMTAP in the purified sample (Figure 2.5, lane 6) was confirmed to be the hexameric form of ApMTAP by MS analysis (paper II, fig. 1B, table S1, S2). But the second band in the purified DgPNP (Figure 2.5, lane 2) has not been tested yet.

To evaluate the enzyme cost for the future process engineering, an estimation of volumetric yield and the productivity per OD_{600} based on the final purified enzymes is provided in Table 2.11. It is shown that the expression levels agree well

with the CAI prediction, which in the order of expression difficulty from high to low was: ApMTAP > TtPyNP > DgPNP > GtPyNP > GtPNP.

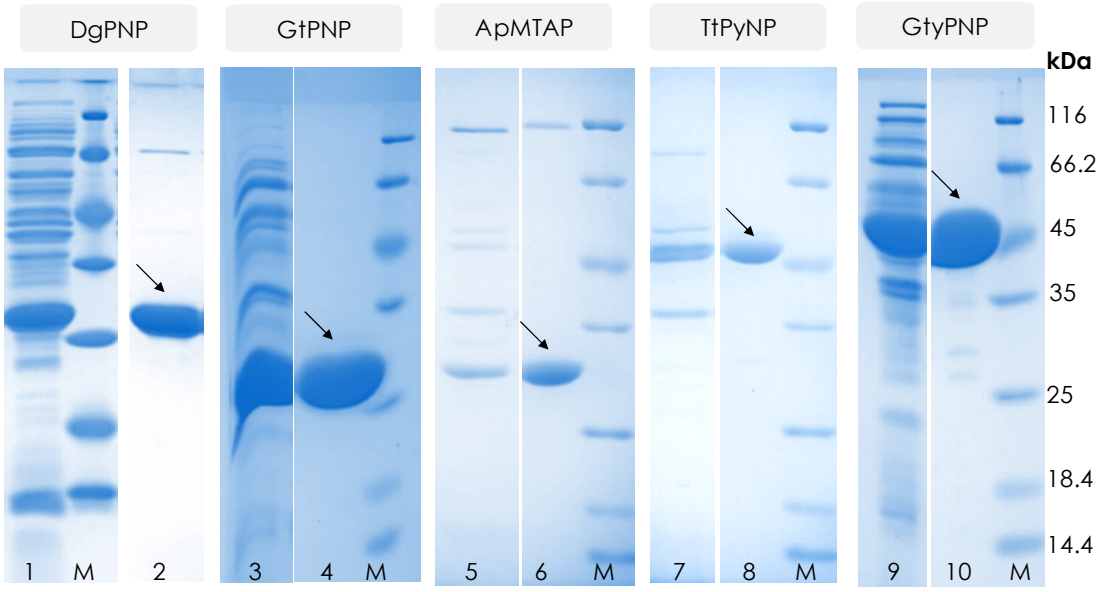


Figure 2.5. SDS-PAGE analysis of recombinant NPs overexpressed in E. coli BL21. Soluble fraction after incubation of the cell lysis at 55 °C (lane1), 65 °C (lane3), 85 °C (lane5), 80 °C (lane7) and 60 °C (lane9) for 15 min. Pooled purified enzymes (lanes 2, 4, 6, 8 and 10) were obtained from the Ni ion affinity chromatography and followed desalting chromatography. M: marker of protein molecular weight.

Table 2.11. Optimized production yield of the purified enzymes.

Parameters	DgPNP	GtPNP	ApMTAP	TtPyNP	GtPyNP
Medium	EnPresso	TB	EnPresso	EnPresso	TB
Cultivation time [a] [h]	36	7.5	36	36	7.5
Volumetric yield [b] [mg L^{-1}]	83.5	297.0	20.8	22.8	166.5
Productivity [mg per OD_{600}]	1.7	7.4	0.5	0.5	6.0

[a] The whole time of the main culture (starts from inoculation).
[b] 400 ~500 mL culture.

2.3. Characterisation

To make use of the thermostable nucleoside phosphorylases, the purified enzymes were characterised in the aspects of the enzymes' temperature optimum, stability, substrate specificity, kinetic studies, pH optimum, and the multiple sequence alignment. These characteristics are the basis for their further immobilization and application for syntheses.

2.3.1. Temperature optimum

The recombinant expression in *E. coli* resulted in moderate to high yields of biologically active enzymes. The effect of temperature on the activity of each enzyme was examined from 30 °C to 99 °C. Except for TtPyNP and ApMTAP, whose reaction rates increased with the temperature until the highest point tested (99 °C), DgPNP, GtPNP, and GtPyNP showed an optimum temperature at 55 °C, 70 °C, and 60 °C, respectively (paper I, fig. 4; paper II, fig. 2).

It is worth to note that for the hyperthermozymes TtPyNP and ApMTAP lower temperatures were unfavourable concerning their activity. For example, the reaction rate of TtPyNP at 60 °C was less than 5 % compared to 99 °C, and 12.7 % compared to 80 °C (paper I, fig. 4). At 40 °C the residual activity of TtPyNP was only 1 % of that at 80 °C (data not published). The situation was similar for ApMTAP: from 30 °C to 40 °C, less than 5 % activity remained compared with its activity at 80 °C (paper II, fig. 2). Compared with the specific activity of GtPyNP at 60 °C towards Urd (51.5 U mg^{-1}, paper II, table 3), TtPyNP at 80 °C, 70 °C and 60 °C showed an activity as 3-fold, 1-fold and one-third of that of GtPyNP, respectively. In other words, when Urd is used as the pentofuranosyl donor (see Scheme 1.1) and the reaction temperature is below 70 °C, it is advantageous to use GtPyNP than TtPyNP because of its higher activity and lower enzyme cost (Table 2.11).

Specifically, for TtPyNP and ApMTAP, the dramatic loss of activity at low temperature could be attributed to the rigidification of their structures which need high energy (temperature) to trigger the conformational changes and to induce the fluctuations of the backbone to the active site (Eisenmesser et al. 2002; Yon et al. 1998). However, it was reported that the hyperthermophilic

enzymes might be activated at the suboptimal temperature by some denaturants (e.g. urea and guanidinum·HCL), detergents (e.g. Triton X-100 and SDS), and solvents (Vieille and Zeikus 2001). It will be of interest to test whether urea or SDS et al. could activate TtPyNP and ApMTAP at low temperature.

2.3.2. Stability

Thermostability

The thermal characteristics of the PyNPs (paper I, fig. 4) and the PNPs (paper II, fig. 3) studied here along with the other reported thermostable PyNPs and PNPs were summarized in paper I (table1) and paper II (table 1), respectively. Thus, the NPs in this study at the optimal or operable temperature between 55 °C and 80 °C were proven to be thermostable for at least 8 h. ApMTAP with an estimated half-life of 69 h at 90 °C was the most thermostable enzyme in this study.

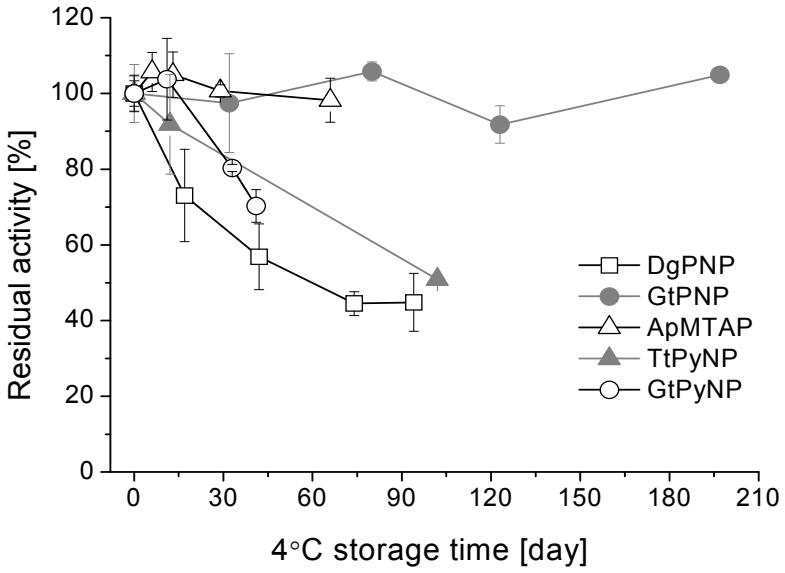


Figure 2.6. Thermostability of purified NPs at 4°C. NPs were stored in 10-50 mM potassium phosphate buffer (pH 7.0) with the concentration of 2~6 mg mL^{-1}. Activity was determined as the standard assay described in the "materials and methods" part in paper I & II. The error bars indicate the standard deviation (n=3).

In view of their storage stability at 4 °C GtPNP and ApMTAP were stable at 4 °C for at least two months without activity loss, while DgPNP, GtPyNP and TtPyNP had obvious activity lost over three weeks (Figure 2.6).

It should be noticed that the above mentioned enzyme thermostability could be considered as the intrinsic thermostability of the enzymes because all the tests were performed in absence of any stabilizing agent (e.g. glycerol, polyols, arginine, or urea). It is conceivable that with the addition of glycerol, for instance, which has been proven to be effective for protein stabilization and prevents protein aggregation (Vagenende et al. 2009), the half-life of the studied enzymes at working temperature as well as at 4 °C would be extended.

The effect of organic solvents

Organic solvents are of great importance in enzymatic processes involving substrates with a low solubility in aqueous media. A suitable organic solvent allows a high reactor substrate concentration and may cause a higher biocatalyst activity, easier product recovery (except for DMSO) and alleviate substrate or product inhibition (Owusu and Cowan 1989). Further, the thermostability of thermozymes is generally considered to be associated with a higher resistance to chemical denaturants (Vieille and Zeikus 2001; Owusu and Cowan 1989). To fully exploit the synthetic capacity of thermostable NPs, the stability of PNPs in the presence of dimethyl sulfoxide (DMSO), MeOH and sodium dodecyl sulphate (SDS) were examined.

The residual activity of DgPNP, GtPNP and ApMTAP all decreased with a higher DMSO concentration (Figure 2.7). Taken GtPNP as an example, its residual activity decreased from 94 % to 59 % when DMSO was increased from 5 % to 20 % (Figure 2.7). GtPNP (0.33 mg mL^{-1}) was incubated with KP buffer (pH 7.0) and 10 % DMSO at 65 °C for 23 h, only 15 % activity remained if KP buffer was 2 mM, while 50 % activity was maintained if KP buffer was 50 mM. It is supposed that the abundant phosphate ion may bind to the enzyme and keep its active site from denaturation.

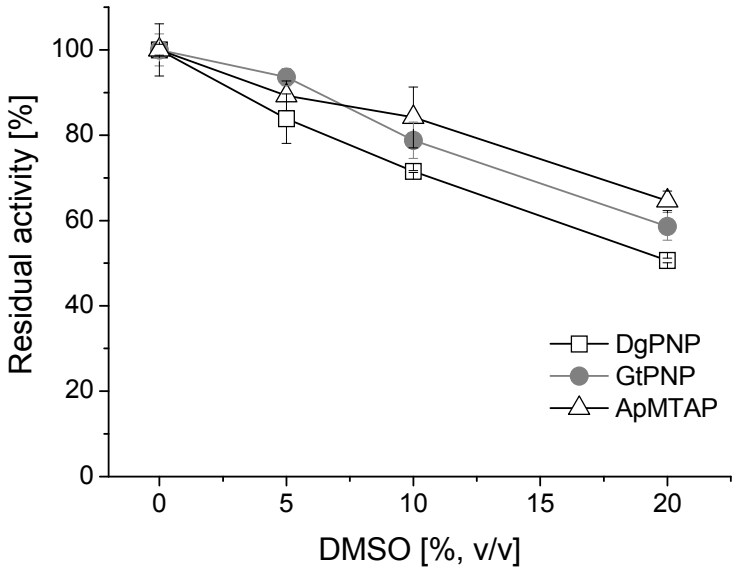


Figure 2.7. Residual activity of PNPs in the presence of DMSO. Reaction conditions: 1 mM Ino in 2 mM Na-phosphate buffer (pH 6.5) containing DMSO 0~20 %, proper amount of enzyme was added to control the reaction conversion <10 % (in 2 min). Reaction temperature: DgPNP 55 °C, GtPNP 70 °C, ApMTAP 80 °C. Samples without DMSO were taken as 100 % residual activity. The error bars indicate the standard deviation (n=2).

To investigate the effect of MeOH on enzyme activity, DgPNP, GtPNP and ApMTAP (final 0.1 mg mL^{-1}) were incubated with 33 % and 50 % MeOH at 30 °C for 20 min individually. No matter 33 % or 50 % MeOH was applied, there was only 1 % activity left for ApMTAP, and 0.6 % activity left for DgPNP and GtPNP. Therefore, 33 % or 50 % MeOH could be used as the agent for stopping the enzymatic reaction.

The stability of PNPs in the presence of 0.02 % SDS (m/v) was tested at 30 °C for 2 h incubation. Unexpectedly, ApMTAP maintained its activity, while DgPNP and GtPNP were completely inactive after incubation (data not shown). The results agree with the SDS-PAGE analysis that hexameric ApMTAP was visible on the gel when DTT was not applied (paper II, fig. 1B).

In summary, all the NPs studied here exhibited expected thermostability according to their optimum temperatures. The PNPs were also examined for their stability in solvents. ApMTAP showed the best resistance to heat denaturation (90 °C) and resisted dissociation even in presence of SDS as well. However, all the tested enzymes lost their activity to some degree in the presence of organic

solvents. It would be advisable to immobilize the enzymes to further improve their stability in organic solvents.

2.3.3. Kinetic properties and substrate specificities

Kinetic properties

To characterise the kinetic properties of the enzymes for answering the questions about (i) how fast can the enzyme catalyses the reaction (k_{cat}), (ii) how tight does the substrate bind to the enzyme (K_m), and (iii) what is the substrate specificity of the enzyme for different substrates (k_{cat}/K_m), we measured the kinetic parameters of k_{cat}, K_m, k_{cat}/K_m and V_{max} for each thermostable NP studied here towards their natural substrates. The results along with the related NPs reported by others were published in paper I (table 2) and paper II (table 2).

Comparing the specificity constants (k_{cat}/K_m), both TtPyNP and GtPyNP are more specific for Urd than for Thd, and the specificity is more pronounced for GtPyNP (paper I). DgPNP, GtPNP and ApMTAP are more specific for Ado than for Ino, which is more pronounced for DgPNP, but it is opposite for EcPNP (paper II).

From a viewpoint of the practical synthesis, the k_{cat} or V_{max} may be the most meaningful parameters because they describe the catalytic efficiency of the enzyme under the practical conditions where the substrate is sufficient. For TtPyNP and GtPyNP, the k_{cat} value of both enzymes is nearly three times higher than that of EcUP towards Urd (280 s^{-1}vs 98 s^{-1}); the k_{cat} value of TtPyNP towards Thd is more than three times higher than that of EcTP (679s^{-1} *vs* 198 s^{-1}) and even eight times higher than that of GtPyNP (679s^{-1}vs 83 s^{-1}). In addition, these kinetic parameters of TtPyNP determined by us agree well with the data obtained from other researchers (Almendros et al. 2012). For PNPs the k_{cat} values of Ino and Ado are all higher than the reported values for the other two hypothermozymes SsMTAP and PfMTAP. Compared with EcPNP, the k_{cat} values of ApMTAP and DgPNP are slightly lower (ApMTAP) or 1~3 times higher (DgPNP) than the k_{cat} values of EcPNP. GtPNP, however, has k_{cat} values (both Ino and Ado) 6 times higher than the values of EcPNP.

It could be supposed that TtPyNP and GtPNP with such high catalytic activities towards natural substrates may retain their kinetic properties towards non-natural substrates as well.

Substrate specificities

To determine the substrate spectrum of the thermostable NPs, a number of nucleosides with different modifications (Figure 2.8) were screened towards the specific activities of the enzymes. The integrated results are summarized in Figure 2.8, which were partially published in paper II (fig. 4 and table 3).

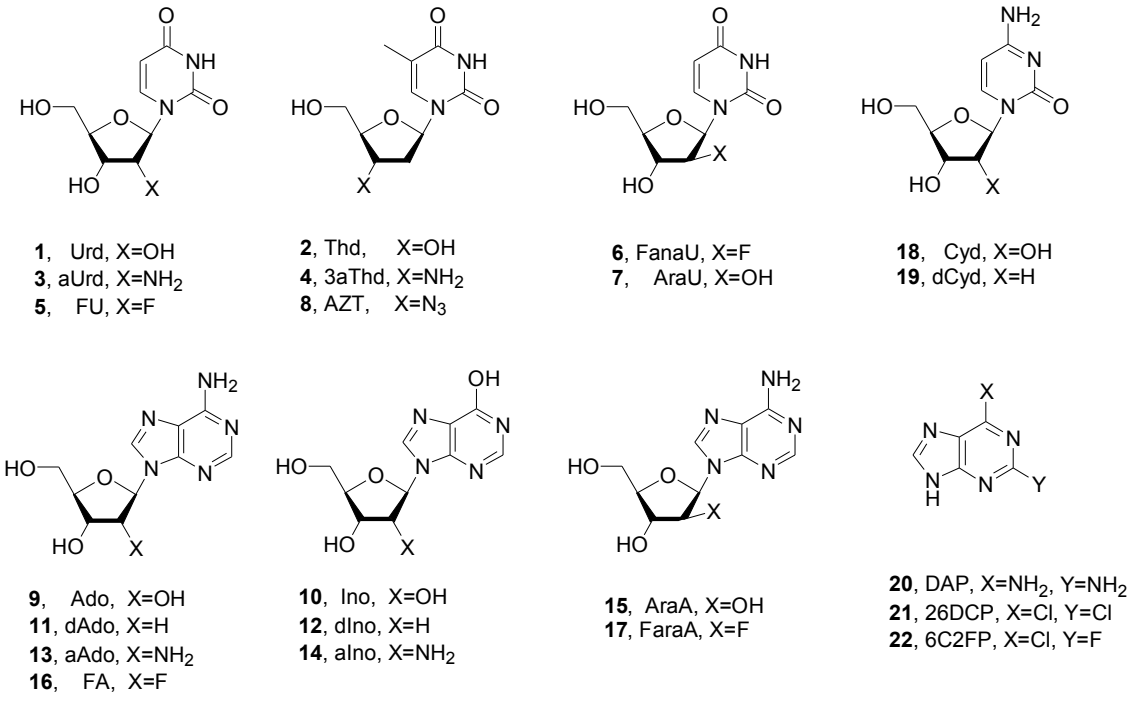


Figure 2.8. Substrates tested for the PyNPs and PNPs in this study.

Considering the transglycosylation, which is catalysed by PyNP and PNP to transfer the sugar moiety from a pyrimidine nucleoside to a purine base, hence the specificity of the PyNP and PNP on the modified sugar moieties (Figure 2.8, **3**-**8** and **13**-**17**) along with the specificity of the PNP on the modified purine bases (Figure 2.8, **20**-**22**) are of great interest.

Table 2.12. Specific activity of PyNPs and PNPs towards natural and nonnatural substrates in this study.

Substrate [a]		Activity of PyNP [U mg^{-1}]		Activity of PNP [U mg^{-1}]		
		TtPyNP 80 °C	GtPyNP 60 °C	DgPNP 55 °C	GtPNP 70 °C	ApMTAP 80 °C
1	Urd	1.54E+2	5.15E+1	9.71E-4	8.32E-4	4.21E-3
2	Thd	2.95E+2	2.24E+1	ND [b]	2.56E-4	6.83E-3
3	aUrd	7.47E+0	1.18E-1	- [c]	-	-
4	3aThd	5.03E+0	8.00E-2	-	-	-
5	FU [d]	5.20E-2	2.00E-3	-	-	-
6	FanaU [d]	5.00E-3	ND[b]	-	-	-
7	AraU [d]	4.60E-2	ND[b]	-	-	-
8	AZT	1.09E-4	ND[b]	-	-	-
9	Ado	8.09E-1	4.58E-1	1.52E+2	3.06E+2	4.13E+1
10	Ino	9.00E-1	3.08E-1	4.37E+1	1.92E+2	4.54E+1
11	dAdo	6.47E-2	2.11E-2	1.92E+2	2.57E+2	1.10E+2
12	dIno	6.67E-2	1.15E-2	1.94E+2	4.98E+2	1.98E+2
13	aAdo	- [c]	-	2.88E-1	1.98E+0	2.52E+0
14	aIno	-	-	4.63E-2	1.97E-1	3.17E+0
15	AraA	-	-	1.66E+0	2.01E+0	1.29E+0
16	FA	-	-	2.29E-2	8.72E-3	1.27E-2
17	FanaA	-	-	2.86E-2	2.08E-2	1.51E-2
18	Cyd	8.67E-4	ND[b]	1.11E-2	1.46E-2	3.68E-2
19	dCyd	-	-	3.55E-2	1.77E-2	1.12E-2
20	DAP[e]	-	-	9.90E+1	1.52E+2	1.06E+1
21	26DCP [e]	-	-	9.10E-1	4.30E+0	3.30E-1
22	6C2FP [e]	-	-	3.30E-1	9.81E+0	2.50E-1

[a] 1 mM substrate in 50 mM phosphate buffer (pH 7.0), if not otherwise stated. See the substrate structures in Figure 2.8 and the chemical names in the "List of abbreviations".
[b] Product peak was not detected.
[c] Not tested.
[d] 1 mM substrate in 10 mM phosphate buffer (pH 7.0).
[e] Apparent activity, see paper IV (Table 3) for more details. 1 mM substrate in 2 mM phosphate buffer (pH 7.0).

The results showed that TtPyNP and GtPyNP do not discriminate pyrimidine nucleosides Urd (**1**) and Thd (**2**), and exhibit a high phosphorolytic activity towards these natural substrates. TtPyNP and GtPyNP also accepted 2'- or 3'-NH_2 substituted pyrimidine nucleosides (aUrd, **3**; and 3aThd, **4**) as substrates, although GtPyNP showed a 63-fold lower activity than TtPyNP towards **3** and **4** (data not

published). Moreover, TtPyNP, but not GtPyNP, was able to phosphorolyse both 2'-deoxyfluroro- *ribo*- and *arabino*- nucleosides (FU, **5**; and FanaU, **6**) as well as arabinofuranosyluracil (AraU, **7**) within 0.005-0.05 U mg^{-1}. Those nucleosides modified at the C2' position of the pentofuranose ring (**3**-**7**, and **13**-**17**as well) are critically important constituents of synthetic oligonucleotides of medicinal potential (Prakash and Bhat 2007; Manoharan 1999). However, it is a great challenge to synthesize this kind of compounds especially for 2'-F (*ribo*- or *arabino*-) (Souleimanian et al. 2012; Watts and Damha 2008) and 2'-NH_2 modified purine nucleosides with traditional chemical methods (Vorbrüggen and Ruh-Pohlenz 2004). The combination of chemical and enzymatic methods is therefore the modern trend of "green chemistry" in the production of valuable nucleosides.

It is interesting to notice that TtPyNP and GtPyNP have considerable activities towards purine nucleosides (**9**-**12**) (paper II). By contrast, PNPs hardly accept pyrimidine nucleosides Urd and Thd, but they accept Cyd and dCyd with the similar activity towards FA (**16**) and FanaA (**17**).

As for the thermostable PNPs, the results proved that they belong to the "high molecular mass PNPs", which have broad specificity and accept 6-oxo (e.g. inosine or guanosine) and 6-amino (e.g. adenosine) purine nucleosides as substrates. Interestingly, the three PNPs are more prone to catalyse 2'-deoxyribosyl nucleosides dAdo (**11**) and dIno (**12**) than the counterpart ribonucleosides Ado (**9**) and Ino (**10**). This behaviour is believed beneficial for the thermophiles to get nutrients from free DNA accidentally in the environments where the thermophiles live (Almendros et al. 2012). Some thermophilic microorganism works as an efficient natural competent machine that allows the internalization of external DNA at a surprising high rate (40 kb s^{-1} $cell^{-1}$). When the internalized DNA is degraded to nucleosides, NPs can further phosphorolyse them into 2'-deoxyribose-1-P and the nucleobase, which can be then used as energy source or is reused for nucleosides synthesis. In this way, thermophiles could survive and grow in a nutrient-poor environment (Almendros et al. 2012; Averhoff and Müller 2010).

All three PNPs can phosphorolyse 2'-fluoro substituted nucleosides FA (**16**) and FanaA (**17**) at a comparable rate to that of TtPyNP towards FU (**5**) and FanaU (**6**),

which meets the basic premise of coupling using PyNP and PNP for the synthesis of 2'-fluoro purine nucleosides.

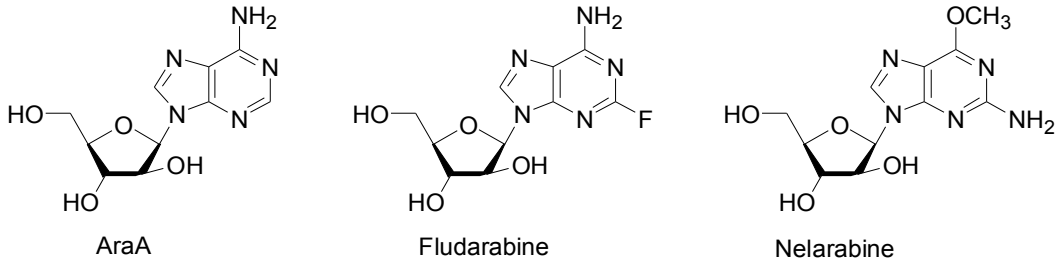


Figure 2.9. Purine arabinosides (AraA, Fludarabine, and Nelarabine).

Purine arabinosides (see Figure 2.9), such as AraA (**15**, vidarabine) – an antiviral drug used in the treatment of many viral diseases (Field and De Clercq 2004), 9-(β-D-arabinofuranosyl)-2-fluoroadenine (Fludarabine) and 2-amino-9-(β-D-arabinofuranosyl)-6-methoxypurine (Nelarabine) – anticancer drugs for B-cell and T-cell lymphocytic leukaemia, respectively (Robak et al. 2006), have gained considerable attention due to their antiviral or antitumor activities. Several strains were reported containing PNP with phosphorolytic activity towards purine arabinosides. Although with low reaction rate, whole cells (Nóbile et al. 2012) or crude extracts (Wei et al. 2008a) were investigated for the synthesis of different purine arabinosides. The thermostable DgPNP, GtPNP and ApMTAP showed activities between 1.3–2.0 U mg^{-1} towards AraA (Table 2.12, not published), which is higher than the reported activities for other PNPs (AhPNPII from *Aeromonas hydrophila*, 0.9 U mg^{-1}; CkPNPII from *Citrobacter koseri*, 0.8 U mg^{-1} (Serra et al. 2013b); and a commercial PNP (Taverna-Porro et al. 2008), enzyme source unknown, 0.003 Umg^{-1})). This suggests their potent biocatalytic properties for the synthesis of purine arabinosides drugs.

Furthermore, the three PNPs are widely tolerant to 2,6-modified purine bases (e.g. DAP, 26DCP and 6C2FP, **20**-**22**), which opens the door for the synthesis of various nucleosides modified at not only sugar moiety but also at the purine base. Among the three PNPs, GtPNP showed superior activities towards the tested natural purine nucleosides (**9**-**12**) and purine bases (**20**-**22**); while DgPNP was the best candidate for phosphorolysis / synthesis of the 2'-fluorinated purine nucleosides (FA and FanaA, **16**-**17**), also because of the lability of FanaU, the

reaction for FanaA synthesis has to be performed at low temperature (paper I, scheme 3). ApMTAP is characterised by a relatively broad substrate spectrum and has a relative high specific activity towards the 2'-amino purine nucleosides aAdo and aIno (**13**-**14**).

In summary, TtPyNP has the following preference for its substrates: Thd > Urd > aUrd > 3aThd > Ino > Ado > dIno > dAdo > FU > AraU > FanaU >> Cyd > AZT. GtPNP showed its substrate preference as following: dIno > Ado > dAdo > Ino > DAP > 6C2FP > 26DCP > AraA > aAdo > aIno > FanaA > dCyd > Cyd > FA >> Urd > Thd.

These results provide a theoretical base for the selection of the enzymes and further applications.

2.3.4. pH optimum

The effect of pH on the reaction rate was investigated to define the proper pH for the optimal enzyme activity and stability.

For PyNP, Urd was tested as a substrate for both, TtPyNP and GtPyNP, at different pH values (50 mM K-phosphate); FU and AraU were tested only for TtPyNP because GtPyNP cannot cleave them. Unexpected, towards different substrates TtPyNP exhibited different pH optima – a pH optimum of 6.8 -7.0 was observed for Urd, but a pH plateau between 7.5~8.5 was appeared in the reactions of FU and AraU (Figure 2.10, closed symbols). GtPyNP activity towards Urd showed a similar trend of the pH-dependency to that of TtPyNP towards FU (Figure 2.10, open squares). In the literature, PyNP from *Geobacillus stearothermophilus* (GsPyNP), showing an 83 % protein sequence identity with GtPyNP, was reported to have a broad pH optimum from 7.0 to 11.5 towards 5-methyluridine (Hori et al. 1990), and an optimal pH of 7.2 towards Urd and Thd (Saunders et al. 1969). TtPyNP, on the other hand, was reported having an optimal pH plateau between 5.0 and 7.0, but the tested substrate was not indicated by the authors (Almendros et al. 2012).

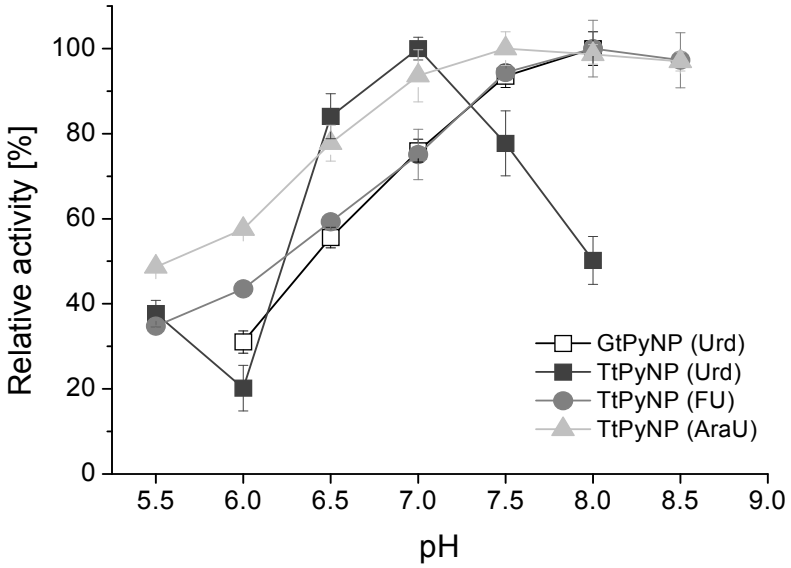


Figure 2.10. Effects of pH on the reaction rate of PyNPs. Reaction conditions: 1 mM substrate in 50 mM K-phosphate buffer (pH 5.5-8.5), diluted enzyme was added into the reaction mixture to control the conversion<5 % (4-10 min). Reaction temperature: TtPyNP80°C, GtPyNP 60 °C. The highest conversion of each substrate and each enzyme was taken as 100 % activity. The error bars indicate the standard deviation (n=3).

Similar contradictory results were also observed for PNPs. PNP from *Geobacillus stearothermophilus* (GsPNPII), showing an 86 % protein sequence identity with GtPNP, was reported with a pH optimum of 7.0~11.0 towards Ino (Hori et al. 1989). We measured the relative pH-dependent activity of DgPNP, GtPNP and ApMTAP with natural and modified substrates (Ado, dAdo, aAdo, Ino, dIno, and aIno). However, the results gave a rather variable picture of activities (paper II, table S3 and fig. S3).

Since proton transfer steps are often directly involved in the enzyme-catalysed processes, pH changes may influence the ionization state of the functional groups in the active site of the enzyme, thus the changes of the reaction rate (should be the maximum velocity when all substrates are saturating) with pH can reflect the protonization or deprotonization of the group which will undergo the rate limiting step (Cleland 1970). Therefore, the diverse pH dependencies on the reaction rate for different substrates implies that different ionization states and/or different groups (mainly on the residues of His, Gln, Asp, Cys, Tyr, and Lys) may be involved in the phosphorolytic reaction depending on the substrates. In spite of the various pH optima for the NPs, pH 7.0 in most cases falls in the pH optimum

plateau and therefore pH 7.0 was used throughout this study if not otherwise stated. However, more detailed studies and experiments are needed (e.g. extending the pH testing range for the PNPs; measuring kinetic parameters under different pH conditions towards the interested substrates) to make a rational interpretation for the results. Thereby, a plot of $\log(V_{max}/K_m)$ *vs* pH will be helpful to reflect the essential ionization groups of the enzyme which is responsible for the binding or catalytic efficiency.

2.3.5. Bioinformatics

In order to understand the relationship between the enzyme structure and its activity as well as the evolutionary relation between the enzymes, bioinformatics tools (protein sequence analysis, phylogenetic analysis, and homology modelling) were used for this purpose. The sequence alignment has been made for the PyNPs (paper I, fig. 1) and PNPs (paper II, fig. 5) separately. Here, the biological data are integrated to get an overview for the NPs on the aspects of their structures and phylogeny, which might provide hints for mining and analysing new enzymes.

The accession numbers of the protein sequence, DNA sequence, and the closest available protein structure (identity is given) for the thermostable NPs which were analysed are listed in Table 2.14.

Multiple sequence alignment

With the development of the bioinformatics tools and the emerging database search methods, the multiple protein sequence alignment based on the structural information of the template has come true e.g. by the server of 3D-Coffee (www.tcoffee.org). It is important to guide the alignment by the protein structure because the sequences may have changed beyond recognition but the similarity of the structure is still detectable by the shape comparisons (Holm and Sander 1996). From this point of view, 3D-Coffee (new name Expresso) runs a BLAST search for the closest homologue of the query sequences within the PDB database, and then uses the hit sequence as the template to guide the alignment, which is significantly more accurate than regular homology-based

methods and the results from it are supposed to give more meaningful information.

PyNP and TP

TtPyNP and GtPyNP were first aligned with PyNP from *G. stearothermophilus* (GsPyNP) and thymidine phosphorylase (TP) from *E. coli* (EcTP) by the ClustalW2 program from the EBI web server (paper I, fig.1). The PyNP active sites according to the literature (Pugmire and Ealick 1998; Mendieta et al. 2004) are highly conserved among them, except that a Met residue of EcTP (Met111) replaces Lys107 of TtPyNP (or Lys108 of GsPyNP). A new sequence alignment by adding *D. geothermalis* PyNP (DgPyNP), *Bacillus subtilis* PyNP (BsPyNP) and *Homo sapiens* TP (HsTP) was made on the 3D-Coffee server. Among them, GsPyNP and BsPyNP have been investigated for the purpose of biocatalytic synthesis (Hori et al. 1990; Hamamoto et al. 1996; Serra et al. 2013a). The alignment result shows again that the Met residue (Met111 in EcTP) is the only change among the known active sites (original results see Appendix 7.1.1; recoloured alignment, see Appendix 7.1.2). This substitution (M111K) appears in EcTP and HsTP, and it may be the key residue which distinguishes ribo- and 2'-deoxyribo- of the nucleoside because it is thought to form a different hydrogen bonding scheme with the phosphate oxygen which binds to the 2'-OH group of the ribose moiety (Pugmire et al. 1998; Pugmire and Ealick 2002). As a result, EcTP is highly specific for 2'-deoxyribose pyrimidine nucleosides (dUrd or Thd), while PyNP doesn't discriminate at the 2'-position of the ribose.

However, the highly conserved active sites between PyNPs cannot explain why TtPyNP has such a high activity towards natural substrates as well as the broad substrate specificity compared to GtPyNP (see Chapter 2.3.3), GsPyNP (Hori et al. 1990; Hamamoto et al. 1996) and BsPyNP (Serra et al. 2013a). Close to the active sites, some substitutions are obvious, for example, Ser83, Ala114, Gly161 and Glu174 in TtPyNP replace the Thr, Gly, Lys and Asp correspondingly in the other PyNPs. Outside of the active site, more variants are observed and DgPyNP seems more close to TtPyNP than to the other. It is true that TtPyNP has a higher sequence identity to DgPyNP (60 %) than to GtPyNP (51 %). Therefore, there is reason to believe that DgPyNP might have as promising catalytic property as

TtPyNP and doesn't need high temperature for the enzyme activation, which is special suitable for the labile reactants, such as FanaU.

PNP and MTAP

A big sequence alignment has been made including high-molecular mass (high-mm) PNP/MTAP and low-molecular mass (low-mm) PNP/MTAP.

From the structure-based alignment (see Appendix 7.1.3) along with the reported active sites details (Bzowska et al. 2000), the differences in the 2D structure of the active sites between high-mm PNP/MTAP and low-mm PNP/MTAP are clearly visible at phosphate-, ribose- and base-binding sites, which are summarized in Table 2.13.

Table 2.13. Comparison of active sites of high-mm PNP and low-mm PNP.

Active site [a]	High-mm PNP (EcPNP/scale No. [b])	Low-mm PNP (HsPNPII/ scale No. [b])
Overall	Phosphate- and ribose- binding sites are buried, while base-binding site is exposed to solvent, and no specific interactions between the enzyme and the purine base are observed	Active sites are on the enzyme surface and exposed to the solvent (hydrophilic environment)
Phosphate – binding site	R24/105 R87/220 S90/223	S33/102 A116/213 Y192/307
Ribose – binding site	VEME cluster 178-181/354-7	VGMS cluster 217-220/354-7
Base – binding site	F159-Y160/313-4 F167/337 I206/384	F200-E201/316-7 AAGG 116-9/223-6 TNKV 242-5/381-4

[a]According to Bzowaska (Bzowska et al. 2000).
[b] Scale number of the alignment indicated above the sequence (see Appendix 7.1.3).

For high-mm PNPs, the phosphate- and ribose-binding sites are buried in the hydrophobic environment while phosphate and ribose are contacted to the charged residues. Taking EcPNP as an example, residues R24, R87 and S90 bind to the phosphate anion and its residue cluster VEME 178-181 binds to the ribose. In contrast, the base-binding of high-mm PNP is mainly due to the non-specific

π–π interactions with aromatic residues (e.g. F159 and Y160 in EcPNP) (Bzowska et al. 2000). Through the alignment, the conserved residues of the enzymes under study (DgPNP, GtPNP and ApMTAP) suggest that they would have a similar substrate specificity as EcPNP, but they could be more tolerant to base while sensitive to sugar moiety due to the binding pattern in their active sites, which was confirmed by our experiments (see Chapter 2.3.3).

For low-mm PNPs, the active sites are exposed to a hydrophilic environment and the phosphate and ribose binding are mainly through hydrogen bonds and hydrophobic interactions. The base binding of low-mm PNPs involves a network of hydrogen bonds between the purine base (N1, N7 and O6) and the side chain of polar or charged residues (e.g. E201, N243 and K244 in HsPNP)(Bzowska et al. 2000). The relative intensive contact of base-binding and relative weak interactions of phosphate and ribose binding imply a stricter specificity for the base moiety and a lower specificity for the pentose (or a quasi-pentose) moiety compared to the high-mm PNPs, which is the case for HsPNP (Stoeckler et al. 1980). According to the alignment, some putative low-mm PNPs, e.g. GtPNPII and ApMTAPII, are supposed to be tolerant to modified sugar moieties and specific to 6-oxopurines.

Hence, the alignment can be used to predict substrate specificity for the new sequence or to guide the site-directed mutagenesis for improving enzyme activity. It is conceivable that novel PNPs with high activity towards both modified base and sugar moiety might come out through protein engineering methods.

UP and PNP

Uridine phosphorylase (UP) accepts uracil and thymine as substrates and does not distinguish ribose and 2'-doxyribose moieties of the pyrimidine nucleosides. Thus, UP has similar substrate specificity as PyNP, unlike PyNP is a dimer, UP reveals a hexameric quaternary structure and its structure is similar to that of hexameric PNP. For instance, the sequence similarity of EcUP and EcPNP is high (with sequence identity of 28 %), which can be seen from the alignment (Appendix 7.1.3).

As UP has broad substrate specificity similar to PyNPs, a thermostable UP (e.g. ApUP) is also considered in this study. Although ApUP was not expressed successfully, it is still worthy to analyse its sequence. A crystal structure of UP from *Streptococcus pyogenes* was recently published (Tran et al. 2011), which shares the highest sequence identity with ApUP (77 %) among the UPs with known structure from PDB. A multiple sequence alignment was made within ApUP, SpUP, EcUP and EcPNP.

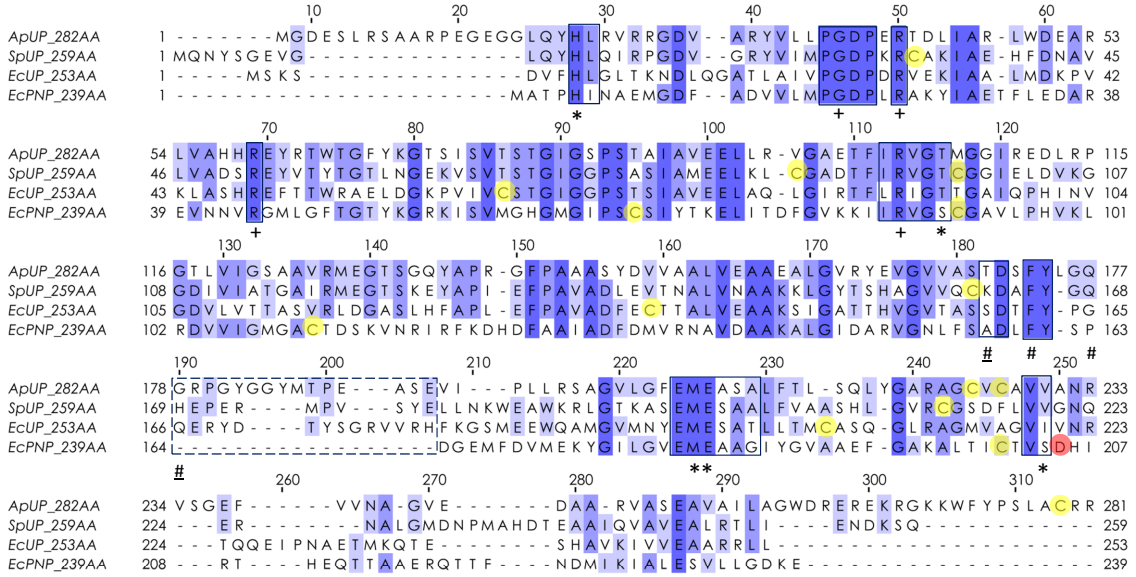


Figure 2.11. Multiple sequence alignment of UP and PNP from A. pernix (ApUP), Streptococcus pyogenes (SpUP), and E. coli (EcUP and EcPNP). The shading represents the degree of sequence identity. The binding sites of SpUP for ribose (*), phosphate (+), base (#) and the novel residues for uracil (#) are indicated below the sequences according to Tran et al. (Tran et al. 2011). Insertion region of UP is revealed in a dashed box. All the Cysteines are highlighted in yellow; one key residue (D204 in EcPNP) for binding to the purine base (Ade) at N7 and at C6-NH_2 is highlighted in red.

The alignment results (Figure 2.11) show that except for the special residues K162 and H169 in SpUP for uracil binding, the other residues in active sites are mainly conserved within EcUP and EcPNP. Residues 166~179 (EcUP numbering) are taken as the UP-specific region (Caradoc-Davies et al. 2004) because it contains the key residues R168 and Q166 for uracil binding and it is not seen in EcPNP sequence. The structural difference between UP and PNP accounts for why UP is specific for pyrimidine but not for purine.

In the active site for base binding, the residues G178, P180 and R233 in ApUP structurally align with Q166, R168 and R223 of EcUP, and H169, P171and Q223 of SpUP, respectively. There is more diverse residues in the base binding site between UPs and PNP. The residues in phosphate and ribose binding site are conserved. Albeit EcUP is about 7-fold more efficient than SpUP (Tran et al. 2011), the similarity and variance of ApUP to EcUP and SpUP make it difficult to predict if ApUP an efficient biocatalyst or not.

ApUP was supposed to contain a disulfide bond, which was attributed to one of the reasons of the unsuccessful protein expression (Szeker 2012). However, the 3D-structure of its homogeneous protein SpUP (PDB: 3QPB) shows that there is no disulfide bond in the hexameric molecule. All the cysteines in the alignment are highlighted (Figure 2.11). In addition, the cysteines of ApUP were predicted to be free cysteine with high score by DiANNA 1.1 web server (Leslie et al. 2002). Therefore, ApUP might not contain disulfide bond; its substrate specificity is similar to EcUP or SpUP.

Phylogenetic analysis

To get an overview of the classification of the studied NPs as well as the other NPs used or not yet used for synthesis, two groups (NP-I and NP-II) of multiple sequence alignments were run on the 3D-Coffee according to their structures. The alignment results were then uploaded to Phylogeny server (www.phylogeny.fr) to produce phylogenetic trees.

According to the phylogenetic tree, NPs consists of two families, NP-I (Figure 2.12 A) and NP-II (Figure 2.12 B), which is also due to the big difference of their structures. In accordance with Pugmire and Ealick (Pugmire and Ealick 2002), the NP-I family is composed of PNP, MTAP and UP and the NP-II family is composed of PyNP and TP.

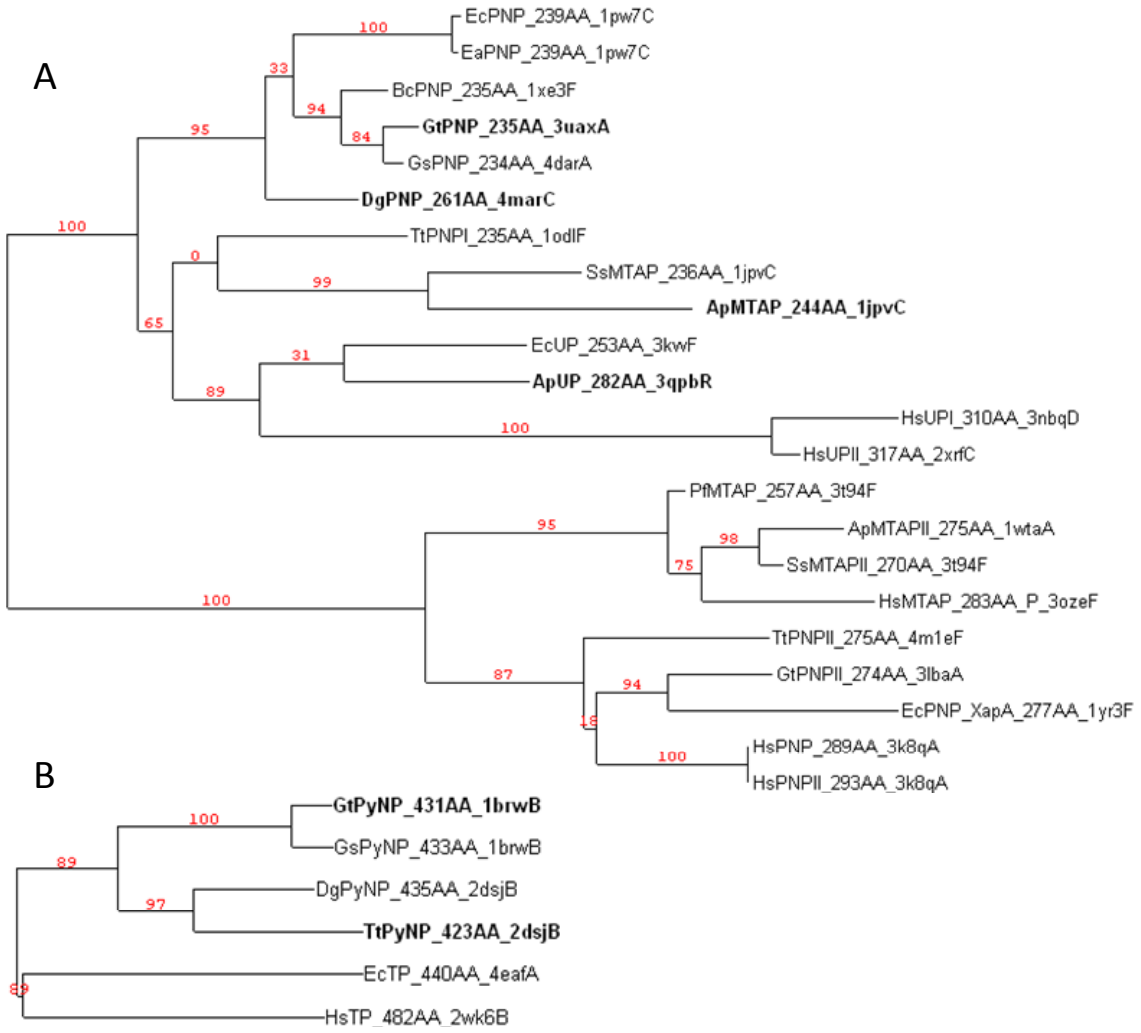


Figure 2.12. Structure-based phylogenetic trees of thermostable (A) NP-I family (PNP, MTAP and UP); (B) NP-II family (PyNP and TP). Human and E. coli NPs are also included as references. The multiple alignments were constructed using 3D-Coffee (Armougom et al. 2006), the trees were produced by PhyML and TreeDyn on the platform of Phylogeny.fr (Dereeper et al. 2008). The branch support values are displayed in red (statistic values), and the branch length indicates the phylogeny distance. For each enzyme, the protein sequence length and the PDB code (structure template used by 3D-Coffee) were indicated following the enzyme name. Enzymes in this study are highlighted in bold. Species abbreviations: Ec= E. coli, Ea= Enterobacter aerogenes, Bc= Bacillus cereus, Gt = G. thermoglucosidasius, Gs = G. stearothermophilus, Dg = D. geothermalis, Tt = T. thermophilus, Ss = Sulfolobus solfataricus, Ap= A. pernix, Hs = Homo sapiens, Pf = Pyrococcus furiosus.

In the NP-I family, high-mm PNPs are generally taken as the PNPs with broad substrate specificity. They accept both 6-oxopurine and 6-aminopurine, while low-mm PNPs or MTAPs are thought to have a strict specificity for 6-oxopurine nucleosides (Bzowska et al. 2000). With new PNPs and MTAPs being characterised

and investigated, it has been found that not all the high-mm PNPs with a similar structure catalysing both, 6-oxopurine and 6-aminopurine, nucleosides. For instance, BcPNP and TtPNPI share high identity with GtPNP (76 % and 41 % respectively) but, unlike GtPNP which accepts adenosine and inosine, BcPNP is highly specific for adenosine (Sgarrella et al. 2007) while TtPNPI is highly specific for inosine (Almendros et al. 2012). Besides, MTAP was classified into the group of trimeric PNPs (or "low-mm PNPs") and with hexameric SsMTAP as an exception (Bzowska et al. 2000; Pugmire and Ealick 2002). Later on, additional hypothermophilic Archaea MTAPs were characterised to be hexamers with a broad substrate specificity towards adenosine, inosine and 5'-deoxy-5'-methlthioadenosine (MTA). SsMTAP, PfMTAP and ApMTAP are also such "exceptions". Strikingly, PfMTAP reveals a high sequence identity with the mammalian-like PNP HsMTAP (52 %) and SsMTAPII (64 %), although it has a distinct substrate specificity (HsMTAP is highly specific for MTA and SsMTAPII is specific for 6-aminopurine) (Cacciapuoti et al. 2011; Cacciapuoti et al. 2007).

The PNPs in the last cluster from the phylogenetic tree A (Figure 2.12 A) (TtPNPII, GtPNPII, EcPNP_XapA, and HsPNPI/II) are "low-mm PNPs" with strict specificity for the purine base. Among them, TtPNPII is a monomer and specific for 6-aminopurine; the others function as a trimer and are specific for 6-oxopurine. It is noteworthy that many bacteria and Archaea contain two PNPs, one with substrate specificity similar to "high-mm PNPs", and another one similar to "low-mm PNP". As discussed before (Table 2.13), two kinds of PNPs have their own strength and weaknesses to adapt modified substrates.

PyNPs and TPs from the NP-II family are easily categorized by the phylogeny tree (Figure 2.12 B). In agreement with the alignment results shown before, DgPyNP displays a close relation with TtPyNP, which might become another promising biocatalyst for nucleosides synthesis.

Overall, through bioinformatics study on the enzymes, lots of structure-active relationships especially for the information about active sites were revealed. In the case of PNPs, their substrate specificity cannot be easily predicted if one only relies on the high sequence identity with a characterised enzyme. The multimeric state, hexameric or trimeric, is not an indicator of a broad or narrow substrate

spectrum for a PNP. Many examples indicated that the key residues in the active site determine the specificity of the enzyme, which is supported by the experiments that switched the enzyme specificity by site-direct mutagenesis of key residues (Stoeckler et al. 1997; Cacciapuoti et al. 2007; Dessanti et al. 2012). Therefore, based on the information of the key residues of the active site, novel enzymes with desired substrate specificity could be generated by rational protein design.

Homology modelling

Among the enzymes studied here, only the crystal structure of TtPyNP is available from PDB. However, the other thermostable NPs can be visualized by homology modelling based on their suitable templates because protein structures are more conserved than protein sequences and the sequence with detectable similarity can be used to produce a structural model of the target (Kaczanowski and Zielenkiewicz 2010).

By BLAST search against the PDB database, the matching structure templates as well as the sequence identity are listed in Table 2.14. The 3D-picture of the enzymes containing their coordinates of all multimer subunits were made by MakeMultimer (http://watcut.uwaterloo.ca/makemultimer) based on the PDB files and shown in Appendix. The structures are coloured by a relative temperature factor (or B-factor), where from blue to red represents the certainty of the position from high to low (Yuan et al. 2005). Hence, a region in pink or red reveals its high flexibility while a blue region reveals a high degree of rigidity in the tertiary structure of the protein.

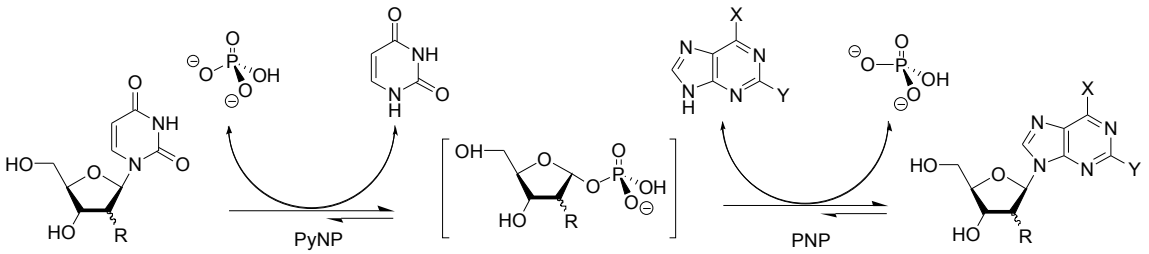
Table 2.14. Accession numbers of the protein sequence and structure.

Enzyme	UniProtKB [a]	GenBank[b]	PDB (enzyme source)[c]	Sequence ident [%][d]
DgPNP	Q1IY92	ABF45792	4MAR (*Meiothermus ruber*)	67
GtPNP	F8CZG4	AEH47728	3UAX (*Bacillus cereus*)	76
ApMTAP	Q9YDC0	NP_147653	1JPV (*Sulfolobus solfataricus*)	45
GtPyNP	F8CV91	AEH47344	1BRW (*G. stearothermophilus*)	83
TtPyNP	Q72HS4	AAS81754	2DSJ (*Thermus thermophilus HB8*)	98
ApUP	Q9YA34	NP_148386	3QPB (*Streptococcus pyogenes*)	77

[a] Protein sequence number for UniProtKB (www.uniprot.org).
[b] Nucleotide sequence number for GenBank (www.ncbi.nlm.nih.gov/genbank).
[c] Protein structure number for PDB (www.rcsb.org/pdb) with the highest amino acid sequence identity to the entry enzyme via a BLAST search against the PDB database.
[d] Protein sequence identity of the entry enzyme to the enzyme with solved structure from PDB (see left).

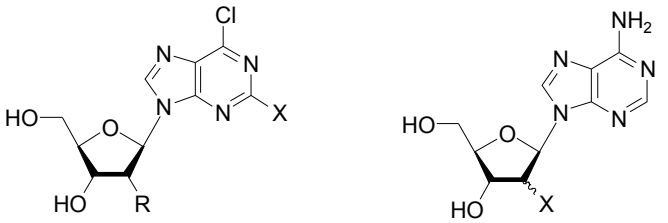
Chapter 3. Transglycosylation processes

In this work, the thermostable nucleoside phosphorylases are investigated as biocatalysts to be utilized for nucleosides synthesis. Particularly, by coupled use of PyNP and PNP, a transglycosylation reaction can be performed efficiently. This includes the transfer of a pentofuranose moiety from an available pyrimidine nucleoside to a purine base. In other words, a substrate pyrimidine nucleoside will be phosphorolized by PyNP into the intermediate α-D-ribose-1-phosphate (R-1-P), and the latter will be connected with a purine base which is the substrate for the PNP, thus a new purine nucleoside is synthesized. The basic synthetic strategy is illustrated in Scheme 3.1.



Scheme 3.1. One-pot enzymatic transglycosylation for the synthesis of purine nucleosides. Short arrows for PyNP and PNP indicate not desired reactions.

There are three reasons for targeting purine nucleosides rather than pyrimidine nucleoside as the desired product: (i) Both PyNP and PNP catalyse reversible reactions, and the equilibrium of the phosphorolysis is usually shifted to the nucleoside formation, however this is more pronounced for the reactions catalysed by PNP than for PyNP (Jensen and Nygaard 1975; Bzowska et al. 2000). Thus the coupled reactions proceed in the direction of purine nucleoside synthesis. (ii) It is effective to use pyrimidine nucleosides as pentofuranose donor because the released pyrimidine base from the first step will not become the competitor of the reactant (purine base) for the second step (Scheme 3.1). (iii) Compared with pyrimidine nucleosides, purine nucleosides are more difficult to synthesize chemically; therefore it is cost-effective to produce purine nucleosides enzymatically.



23, 26DCP-R, X=Cl, R=OH
24, 6C2FP-R, X= F, R=OH
25, 26DCP-dR, X=Cl, R=H
26, 6C2FP-dR, X= F, R=H

9, Ado, X=OH
16, FA, X=*ribo*-F
17, FanaA, X=*arabino*-F

Figure 3.1. Products from the enzymatic transglycosylation reactions.

In our case, the desired products are halogenated purine nucleosides (Figure 3.1) either with the purine base modified at 2,6-positions (2,6-dichloropurine and 2-chloro-6-fluoropurine) or with the sugar moiety modified at the 2'-position (2'-dexoy-2'-fluoro *–ribo* and *–arabino*). The 2,6-dihalogenated purine nucleosides are of great interest as precursors for the synthesis of numerous purine nucleosides with therapeutic potential. More details are discussed in paper IV. Nucleosides modified at the 2'-position are vital important constituents of artificial oligonucleotides for the RNA interfering therapeutics (Prakash and Bhat 2007; Manoharan 1999). Among them, 2'-deoxy-2'-fluoro modifications are considered as the best-tolerated alteration with respect to maintain silencing activity, however, 2'-fluoro-substitued purines were commercially unavailable before 2007 and only pyrimidine-modified oligonucleotides were attempted (Blidner et al. 2007), probably due to the great difficulties of their preparation. In addition, purine 2'-dexoy-2'-fluororibosides were shown to have a very potent antiviral activity (Tuttle et al. 1993). Also, 2'-fluoroarabino nucleic acid has been turned out as the best example of chemically modified single-stranded antisense oligonucleotides (Watts and Damha 2008; Watts and Corey 2012). Therefore, this work focuses on the enzymatic transglycosylation with the aim of developing a more efficient and "green" method to produce aforementioned purine nucleosides.

Fortunately, the previous results of the enzyme characterisation (Chapter 2.3) demonstrated that TtPyNP is able to catalyse 2'-fluoro substituted uridine and all PNPs (DgPNP, GtPNP and ApMTAP) recognize 2'-modified pentofuranose as well

as 2,6-dihalogenated purines as substrates. Hence, this offers the possibility of developing a practical synthesis of interesting purine nucleosides.

In one-pot enzymatic transglycosylation (Scheme 3.1) the reaction temperature depends on the enzyme with the lower optimal temperature. Thus, if GtPNP (optimal temp. =70 °C) is used together with TtPyNP (optimal temp. > 90 °C), the upper limit of the reaction temperature is 70 °C; if GtPNP is combined with GtPyNP (optimal temp. = 60-65 °C), 65 °C is the highest temperature to be used.

Based on the substrate specificity of the used enzymes described in Chapter 2.3, at 65-70 °C TtPyNP and GtPyNP are interchangeable because of their similar phosphorolysis activity towards uridine and thymidine. However, when the pentofuranose donor is 2'-fluoro substituted uridine (FU or FanaU), TtPyNP must be used since GtPyNP shows no activity towards FU and FanaU; while TtPyNP and the three PNPs accept 2'-fluoronated nucleosides as substrate.

According to the product yields and the reaction equilibrium states (monitored by HPLC, paper II), the transglycosylation processes studied here are divided into three types which are correlated with three different pentofuranosyl donors: uridine (Urd, **1**), thymidine (Thd, **2**) and 2'-fluoro substituted uridines (FU, **5**; and FanaU, **6**). For each process type, the experiment results and discussion will be covered in details as below.

3.1. Type1: Equilibrium-controlled reactions

3.1.1. One-pot transglycosylation

When uridine was used as the pentofuranosyl donor, and adenine (Ade), 2,6-dichloropurine (26DCP, **21**) or 2-chloro-6-fluoropurine (6C2FP, **22**) were applied as purine base acceptors, the one-pot transglycosylation showed a similar progress curve for all three reactions; the reaction reached an equilibrium at some time point and the product concentration did not change over the reaction time. With different enzyme combinations and even at different temperatures (GtPyNP &DgPNP at 55 °C; GtPyNP & GtPNP at 65 °C; TtPyNP & ApMTAP at 80 °C) very similar product yields were obtained for the respective substrate. However, when different purine bases Ade, 26DCP or 6C2FP served as substrate, different product yields were obtained: 88 % for adenosine (Ado, **9**), 56 % for 2,6-

dichloropurine riboside (26DCP-R, **23**) and 67 % for 2-chloro-6-fluoropurine riboside (6C2FP-R, **24**).

The reaction process curves of 26DCP-R and 6C2FP-R are given in Figure 3.2. It can be seen that the equilibrium is established within 30 min, and then the concentration of the product and the side-product (uracil) keep constant.

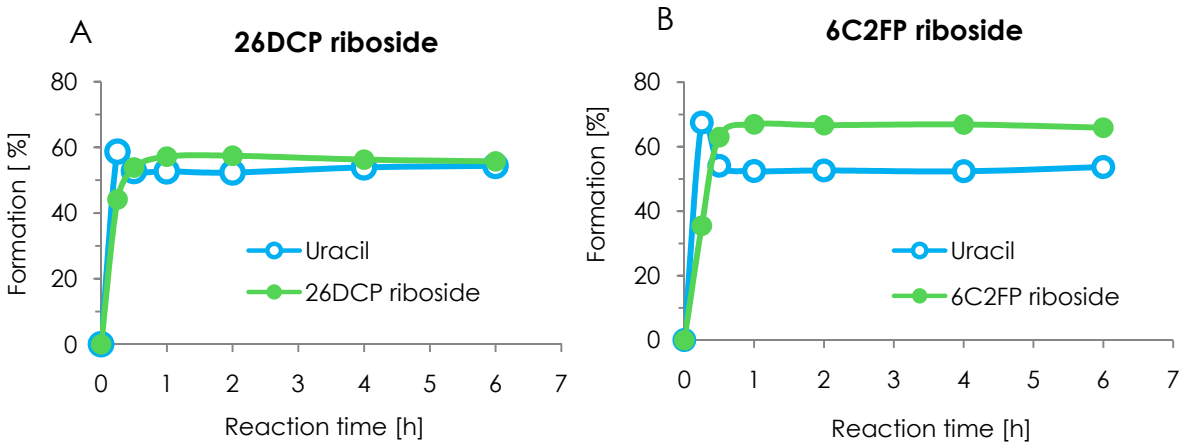


Figure 3.2. Progress curves of enzymatic synthesis of (A) 26DCP riboside and (B) 6C2FP riboside at 65 °C by GtPyNP and GtPNP (100 μg mL^{-1} each). Substrate solution: 2 mM uridine and 1 mM purine base (26DCP or 6C2FP) in 2 mM Na-phosphate buffer (pH 6.5). Formation ofuracil and product were calculated from [Uracil]/[Urd]$_{initial}$×100 % and [product]/[purine base]$_{initial}$×100 %, respectively.

In the case of 26DCP-R synthesis the product formation, i.e. the equilibrium yield, raised from 57 % to 83 % (Figure 3.3 A) when the substrate ratio of Urd : purine base was changed between 2:1 to 10:1. However, a high substrate ratio (10:1) may be unfavourable for separation and purification of the final product. On the other hand, changing the enzyme amount ratio of GtPyNP: GtPNP from 1:1 to 1: 20 (unit ratio changing from 13:1 to 1:1.5) had almost no influence on the equilibrium yield for the product (~60 %) (Figure 3.3 B, curves 2-4). Only if GtPyNP was applied at a very low concentration of 1 μg mL^{-1} (the enzyme unit ratio of GtPyNP : GtPNP became 1:7) product formation was hindered by the lack of GtPyNP. Consequently the equilibrium yield decreased to 26 %. Interestingly, the phosphorolysis of Urd (i.e. uracil formation) was enhanced from 35 % to 52 % (Figure 3.3 B, chart 1 *vs* 2) when GtPyNP was used alone compared to the coupled use of GtPyNP and GtPNP.

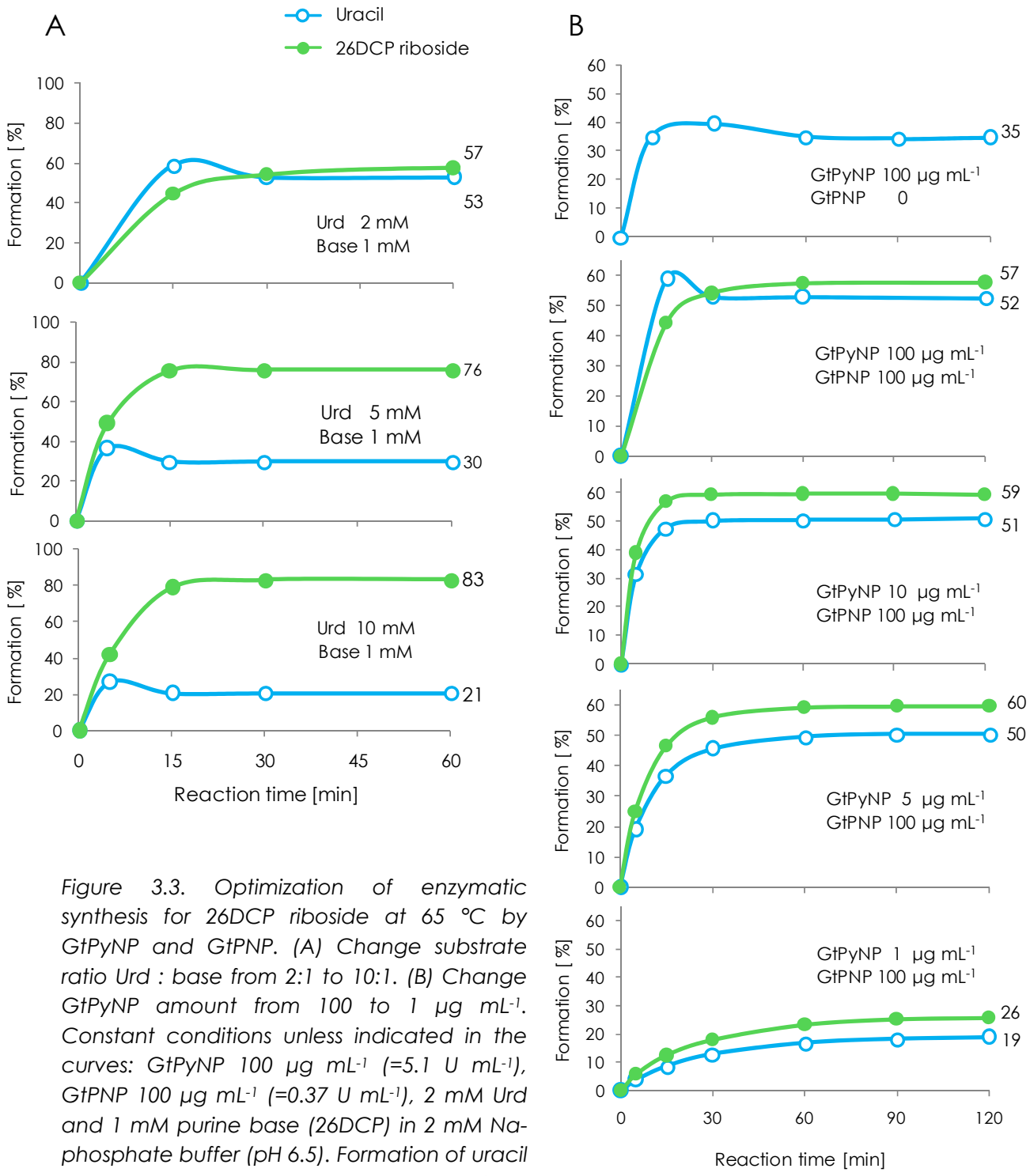


Figure 3.3. Optimization of enzymatic synthesis for 26DCP riboside at 65 °C by GtPyNP and GtPNP. (A) Change substrate ratio Urd : base from 2:1 to 10:1. (B) Change GtPyNP amount from 100 to 1 µg mL^{-1}. Constant conditions unless indicated in the curves: GtPyNP 100 µg mL^{-1} (=5.1 U mL^{-1}), GtPNP 100 µg mL^{-1} (=0.37 U mL^{-1}), 2 mM Urd and 1 mM purine base (26DCP) in 2 mM Na-phosphate buffer (pH 6.5). Formation of uracil and 26DCP-R were calculated from [Uracil]/[Urd]$_{initial}$×100 % and [26DCP-R]/[26DCP]$_{initial}$ ×100 %, respectively.

It is very likely that a similar trend would appear also for the synthesis of 6C2FP-R by changing of substrate ratio or enzyme ratio.

To sum up, the enzymatic syntheses of 26DCP-R (**23**) and 6C2FP-R (**24**) are equilibrium controlled reactions, whereby the equilibrium yield is independent of enzyme combination and the enzyme amount over a wide concentration range (if only their activity is in the same range); but the equilibrium yield depends on the initial substrate concentration.

3.1.2. Single step of the transglycosylation

Urd + Pi

In the practical synthesis, it is favourable to apply the substrate at a concentration as high as possible (if the substrate inhibition is not obvious) for a high volumetric yield and to employ the enzyme as little as possible to reduce the cost. Therefore, it is necessary to examine the influence of excess substrate on the enzymatic reaction.

Urd is the substrate to be applied at a high concentration due to its low cost and high solubility. Urd phosphorolysis was performed under low concentrations of GtPyNP at 60 °C (Figure 3.4) to mimic the condition of excess Urd (i.e. high Urd concentration). The process curves show that with a substrate (Urd 5 μmol mL^{-1}) to enzyme (GtPyNP 1.2 and 0.0062 U mL^{-1}) ratio of 4.2 and 806, the initial velocity of the reaction (in 1 min) is 32 and 102 U mg^{-1}, respectively. Hence, a high substrate to enzyme ratio (806) increased the initial velocity of the reaction by a factor of three. In addition, with 0.12 μg mL^{-1} GtPyNP (0.0062 U mL^{-1}) Ura formation was linear with the time (R^2=0.999) during the tested 25 min and the average velocity of the reaction was 91 μmol min^{-1} mg^{-1}, which is only half of the V_{max} value (174 μmol min^{-1}, 60 °C, unpublished data), but the latter was performed with 10 times higher phosphate concentration (50 mM *vs* 5 mM). As a conclusion, no obvious substrate inhibition was observed even at a high concentration ratio of substrate to enzyme.

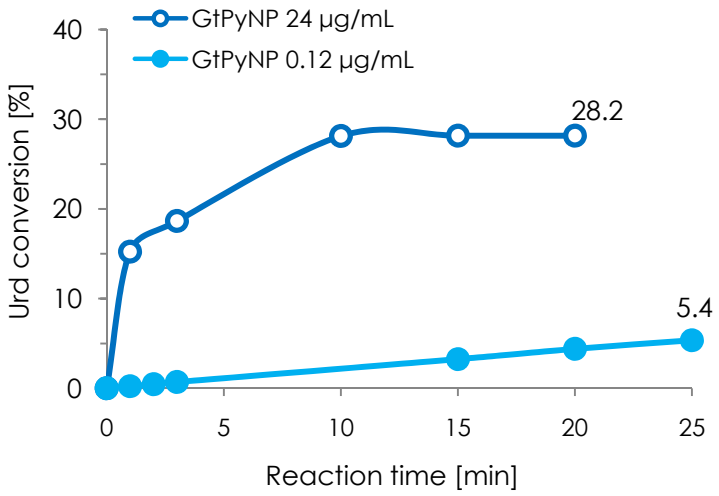


Figure 3.4. Phosphorolysis of Urd by GtPyNP at 60 °C. Reaction conditions: GtPyNP 24 or 0.12 µg mL^{-1} (=1.2 or 0.0062 U mL^{-1}), 5 mM Urd in 5 mM K-phosphate buffer (pH 7.0). Urd conversion was calculated from [Uracil]/[Urd]$_{initial}$×100 %. The highest error bar (not shown) is less than 2 % (n=2).

R-1-P + Ura

Although it is known that the equilibrium of nucleoside phosphorolysis is thermodynamically in favour of nucleoside synthesis (Bzowska et al. 2000), the equilibrium constant (K_{eq}) or the kinetic parameters for the reverse reaction (nucleoside synthesis) has not been described in the literature for PyNPs. This might be due to the very high cost of α-*D*-(2-deoxy)-ribose-1-phosphate ((d)R-1-P) and the instability of (d)R-1-P, as well as due to the challenge for performing the assay for the reverse reaction, while these kinetic data would be important for enzymatic process modelling and engineering. Albeit the equilibrium concentration of R-1-P can be deduced from the concentration of the other components which are measured by HPLC, it is still necessary to attain the experimental data as the proof.

Specifically, the calculation of [R-1-P] is based on the species balances. Once the equilibrium is established, for the reaction with Urd as pentofuranosyl donor,

$$[Urd]_0=[Urd]_{eq} + [Product]_{eq}+ [R\text{-}1\text{-}P]_{eq},$$

also $$[Urd]_0=[Urd]_{eq}+ [Ura]_{eq},$$

hence $[R\text{-}1\text{-}P]_{eq} = [Product]_{eq} - [Ura]_{eq}$. Eq. 1

Likewise, $[Pi]_0=[R\text{-}1\text{-}P]_{eq}+ [Pi]_{eq}$,

thus $[Pi]_{eq}=[Pi]_0 - [R\text{-}1\text{-}P]_{eq}$. Eq. 2

Where "$[\ ]_{eq}$" means equilibrium concentration and "$[\ ]_0$" means initial concentration, while "Pi" stands for phosphate and Ura stands for uracil.

The transglycosylation (Scheme 3.1) can be divided into two steps:

Step1 Urd + Pi ↔ R-1-P + Ura, where

$$K_{eq1} = \frac{[\mathrm{R-1-P}]_{\mathrm{eq}} \times [\mathrm{Ura}]_{\mathrm{eq}}}{[\mathrm{Urd}]_{\mathrm{eq}} \times [\mathrm{Pi}]_{\mathrm{eq}}}$$

Step2 R-1-P + Base ↔ Product + Pi, where

$$K_{eq2} = \frac{[\mathrm{Product}]_{\mathrm{eq}} \times [\mathrm{Pi}]_{\mathrm{eq}}}{[\mathrm{R-1-P}]_{\mathrm{eq}} \times [\mathrm{Base}]_{\mathrm{eq}}}$$

Using Eq. 1 and Eq. 2 along with the other component concentrations measured by HPLC, K_{eq1} and K_{eq2} are calculable. It is worthy to point out that the forward reaction of Step1 is phosphorolysis while Step2 is synthesis. Thus, higher values of K_{eq1} and K_{eq2} are favourable for purine nucleoside i.e. product formation.

As for the overall reaction: Urd + Base ↔ Product + Ura, we have

$$K_{eq} = \frac{[\mathrm{Product}]_{\mathrm{eq}} \times [\mathrm{Ura}]_{\mathrm{eq}}}{[\mathrm{Urd}]_{\mathrm{eq}} \times [\mathrm{Base}]_{\mathrm{eq}}}$$

Theoretically, by HPLC measurements followed by simple calculations, the equilibrium constant of the sub-step reaction and the overall reaction related to the transglycosylation can be obtained.

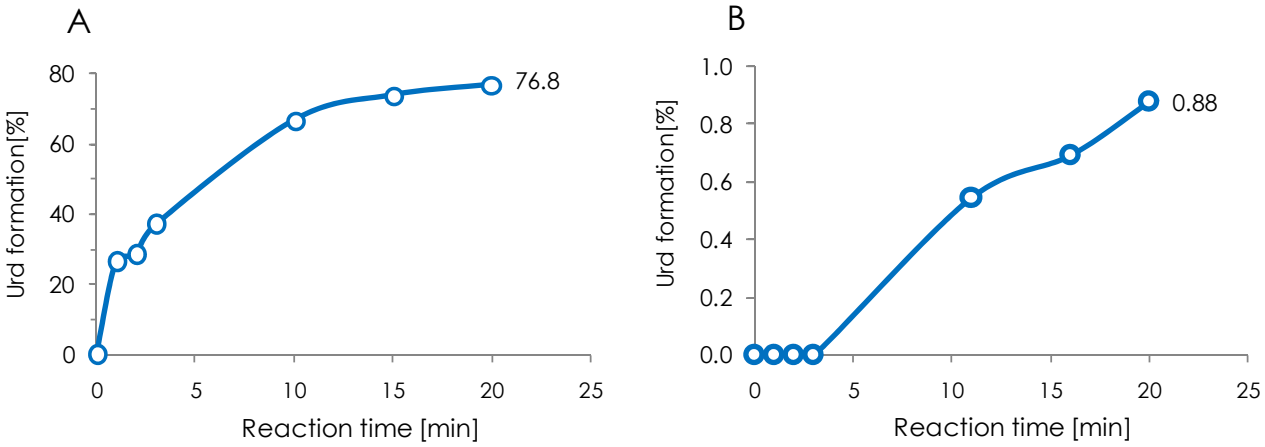


Figure 3.5. Synthesis of Urd from R-1-P and uracil by GtPyNP (A) 24 μg mL^{-1} (e.g. 1.2 UmL^{-1}) and (B) 0.12 μg mL^{-1} (e.g. 0.0062 U mL^{-1}) at 60 °C. Reaction conditions: R-1-P 0.5 mM and uracil 1 mM in water. Urd formation was calculated from [Urd]/[R-1-P]$_{initial}$×100 %. The highest error bar (not shown) is less than 6 % (n=2).

The reverse reaction of Urd phosphorolysis was performed in water with 0.5 mM R-1-P (kindly donated by Sigma-Aldrich) and 1 mM uracil at 60 °C by GtPyNP (Figure 3.5). Compared with the Urd phosphorolysis discussed above (Figure 3.4), with a higher enzyme amount of 24 μg mL^{-1}, for both reactions the apparent reaction rates are much higher than the counterparts with 200-time diluted GtPyNP (0.12 μg mL^{-1}), but the ratio of enzyme specific activity (U per mg enzyme) of high amount enzyme *vs* low amount enzyme for Urd phosphorylysis is 1 : 3 (32 U mg^{-1}: 102 U mg^{-1}; in initial phase) and for Urd synthesis is 2 : 1 (5.5 U mg^{-1} : 2.1 U mg^{-1}; in initial phase). These data indicate that when the substrates are supplied far in excess for GtPyNP, Urd (and phosphate) cause no substrate inhibition but Ura (and R-1-P) may inhibit the enzyme at a slight degree. Extra well-designed inhibition experiments are needed to draw a clear conclusion.

According to the calculation method for the equilibrium constant, based on the data of Urd phosphorolysis (Figure 3.4 A) at 60 °C, in phosphate buffer with a pH of 7.0, $[Urd]_0=[Pi]_0=5$ mM, K_{eq1} was calculated as 0.15, which agrees well with the other experiments performed at the same conditions of temperature and buffer, while K_{eq1} changes independence on the starting point (e.g. $K_{eq1}= 0.17$ for $[Urd]_0= [Pi]_0= 2$ mM; $K_{eq1}= 0.20$ for $[Urd]_0= 5$ mM and $[Pi]_0= 2$ mM). For the reverse

reaction (Urd synthesis), $[R\text{-}1\text{-}P]_0$= 0.5 mM and $[Ura]_0$ = 1 mM, however, this reaction was only close to but not reached the equilibrium (Figure 3.5 A). An estimated K_{eq1} (0.48) was calculated from the last time point with a yield of 78.6 % (Urd formation). Since K_{eq1} should be constant and independent of the starting reaction, if K_{eq1} was assumed to be 0.17, the Urd formation yield would become 88 %.

For Ado synthesis (in the transglycosylation reaction), $K_{eq\ (Ado)} = K_{eq2} = 30$; for Urd synthesis (in the same transglycosylation reaction) $K_{eq\ (Urd)} = (K_{eq1})^{-1} = (0.44)^{-1} = 2.3$; while for Urd synthesis by PyNP alone $K_{eq\ (Urd)} = (K_{eq1})^{-1} = (0.15)^{-1} = 6.7$. These data prove that the equilibrium of the reaction catalysed by PyNP or PNP alone is shifted to the synthesis direction, and this is more pronounced for the formation of purine nucleosides than for the formation of pyrimidines. By coupled use of PyNP and PNP the extent of Urd synthesis is lower than that catalysed by PyNP alone, which is beneficial for product formation. As for the modified substrates, the equilibrium constant will be different, but the overall reaction could progress into the desired reaction.

R-1-P + 6C2FP

The target purine nucleoside can be also synthesized in a single step by PNP alone, but the expensive R-1-P becomes the limitation of this method. Here the work focuses not on the synthesis by this method, but on the reaction mechanism of this sub-step reaction and the influence factors involved, in order to understand the whole transglycosylation process for future control and optimization in practical applications.

In the present study 6C2FP-R (24) was applied as the target product; R-1-P and 6C2FP (22) were used as substrates for GtPNP. In the substrate stability test analysed by HPLC, 2-3 unknown peaks became evident with an increasing reaction temperature (especially above 70 °C) over 20 hours (data not shown). Thus, lower reaction temperatures were adopted with two concentrations of GtPNP for 6C2FP-R synthesis.

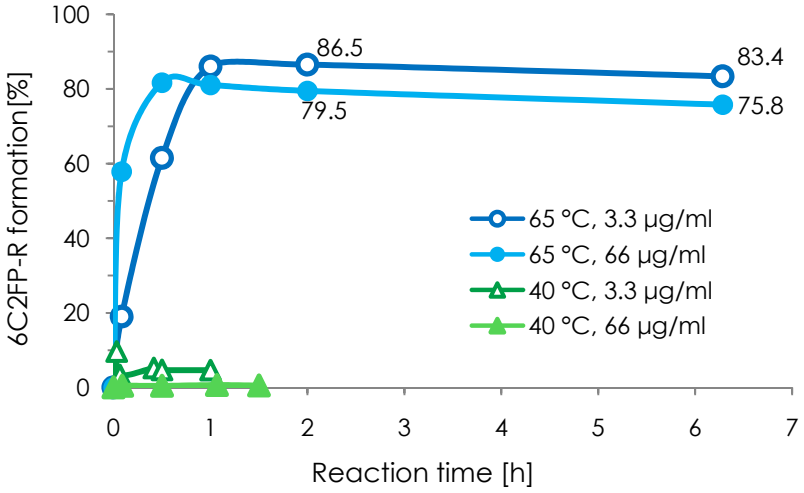


Figure 3.6. Synthesis of 6C2FP-R from R-1-P and 6C2FP at 40 °C and 65 °C by GtPNP. Reaction conditions: GtPNP 3.3 and 66 µg mL^{-1}, R-1-P 0.5 mM, 6C2FP 1 mM in water. 6C2FP-R formation was calculated from [6C2FP-R]/[R-1-P]$_{initial}$×100 %.

Process dynamics are displayed in Figure 3.6. The curves show that the reactions at 65 °C reach the equilibrium in less than one hour. The reaction with low enzyme concentration (3.3 µg mL^{-1}) even obtained a higher product yield (>83 %). Likewise, reactions at 50 °C and 60 °C also showed similar process curves (data not shown).

Unexpectedly, reactions at 40 °C did not work. This could be attributed to fact that the thermozyme GtPNP needs a higher temperature for activation. According to the apparent activity of GtPNP towards 6C2FP (paper IV, table 3), and the temperature dependence of GtPNP activity (towards inosine, paper II, fig. 2) for 6C2FP, 3.3 and 66 µg mL^{-1} GtPNP equal to 0.0071 and 0.14 U mL^{-1} at 40 °C, 0.029 and 0.57 U mL^{-1} at 60 °C, respectively. Obviously, this estimation was not proper, because the reaction with 66 µg mL^{-1} GtPNP at 40 °C (0.14 U mL^{-1}) supposed to be more efficient than the reaction with 3.3 µg mL^{-1} GtPNP at 60 °C (0.029 U mL^{-1}). This implies (i) the temperature-active curves of the thermostable NPs might be varying towards different substrates (especially for modified ones); and (ii) different substrates may have different activation temperature for the enzymes.

In summary, the processes of the equilibrium-controlled reactions were analysed for the whole transglycosylation as well as in the single step reactions. A method

for calculating equilibrium constants of whole and sub-step reactions was proposed, which was confirmed by the experimental data. For the synthesis of dihalogenated purine nucleosides (**25** and **26**), the product yields are limited by the equilibrium conditions. Changing substrate ratios can improve the final yield. 7-Methylinosine (also valid for 7-methyladenosine or 7-methylguanosine) iodide, is a substrate for various PNPs and undergoes non-reversible phosphorolysis (Bzowska et al. 2000). Thus these substances can be used as a pentofuranosyl donor for the second nucleoside synthesis by providing the base and PNP as the only enzyme. It has been reported that the phosphorolysis of 7-methylguanosine by PNP from *Bacillus subtilis* and *Aeromonas hydrophila* at room temperature in 50 mM phosphate buffer (pH 7.5) reached a 100 % conversion (Ubiali et al. 2012). The complete phosphorolysis is useful to shift the equilibrium to the product-generating direction and then to improve the product yield.

3.2. Type2: Reactions which involve unstable intermediates

3.2.1. One-pot transglycosylation

When thymidine (Thd, **2**) served as the pentofuranosyl donor, and 2,6-dichloropurine (26DCP, **21**) or 2-chloro-6-fluoropurine (6C2FP, **22**) as the purine base acceptor, the corrosponding products are 2,6-dichloropurine-2'-deoxyriboside (26DCP-dR, **25**) and 6-chloro-2-fluoropurine-2'-deoxyriboside (6C2FP-dR, **26**), respectively.

The progress curves of the one-pot transglycosylation of **25** and **26** were very unexpected (see Figure 3.7, **25** and **26** were alike in their progress curves, of which only **26** is shown). A maximum product yield is obtained in a short time and then the yield dropped dramatically. This was more evident at high temperature where the product yield decreased to 7 % within 2 h, while at 65 °C the product yield still remained at 17 % after 6 h. On the other hand, Thd phosphorolysis continued proceeding and Thd was almost completely converted into thymine (97 % conversion) at 80 °C after 2 h.

Unlike uridine, thymidine is unstable and its glycoside bond breaks by addition of trichloroacetic acid (TCA, used for stopping the enzyme reaction). Hence, all the reactions which involved thymidine were stopped by MeOH or cold water and

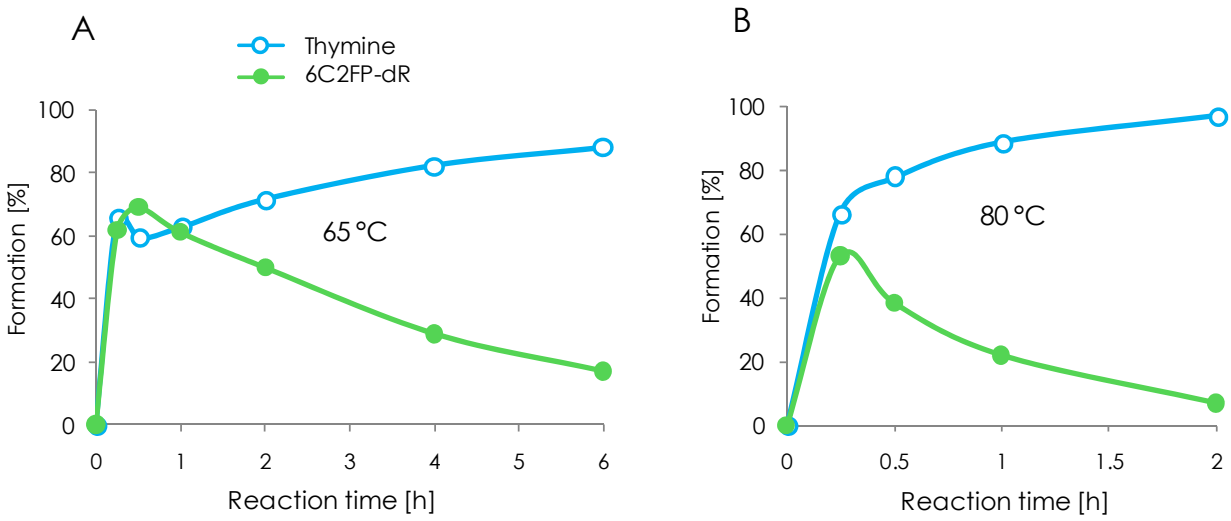


Figure 3.7. Progress curves of enzymatic synthesis of 6C2FP-dR. (A) 65 °C by GtPyNP and GtPNP(0.1 mg mL^{-1} each); (B) 80 °C by TtPyNP and ApMTAP (0.1 mg mL^{-1} each). Substrate solution: 2 mM Thd and 1 mM 6C2FP in 2 mM Na-phosphate buffer (pH 6.5). Formation of thymine and 6C2FP-dR were calculated from [Thymine]/[Thd]$_{initial}$×100 % and [6C2FP-dR]/[6C2FP]$_{initial}$×100 %, respectively.

negative controls (without enzyme) were performed in parallel. The lability of thymidine promts us to assume that either the product or the intermediate α-*D*-2'-deoxyribose-1-phosphate (dR-1-P) is instable and hydrolysis might occur during the reaction process (Szeker 2012). To confirm this hypothesis, single step reactions were performed to disclose the reason.

3.2.2. Single step of the transglycosylation

Single steps of Thd phosphorolysis and 6C2FP-dR synthesis were carried out separately. If dR-1-P is unstable and undergoes hydrolysis, for Thd phosphorolysis, Thd will be continually converted to thymine and no equilibrium should be observed; for 6C2FP-dR synthesis, 6C2FP-dR will be first formed, and is then converted back to 6C2FP because of the elimination of dR-1-P; therefore the yield of 6C2FP-dR (formation) will first go up and then decline.

The process curves of the single step reactions proved this presumption. In the reactions of Thd phosphorolysis, Thd conversion reached 35 % in less than 10 min and then fluctuations of the conversion from 5 to 40 min were observed in the reactions of 40 °C, 50 °C and 60 °C (data not shown). Afterwards, Thd conversion

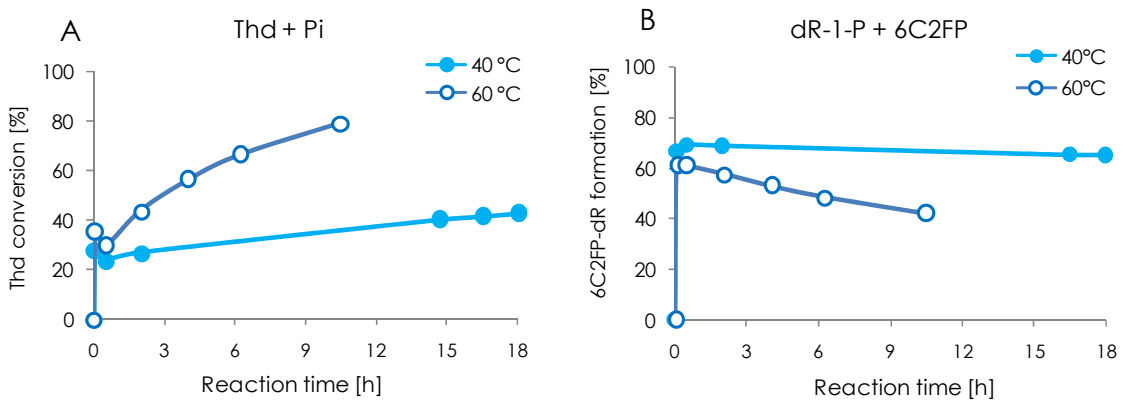


Figure 3.8. Process curves of (A) Thd phosphorolysis by GtPyNP and (B) 6C2FP-dR synthesis by GtPNP. Reaction conditions: (A) 2 mM Thd in 2 mM K-phosphate buffer (pH 7.0), GtPyNP 47.6 μg mL^{-1} (1.1 U mL^{-1} at 60 °C towards Thd). (B) R-1-P 0.5 mM, 6C2FP1 mM in water, GtPNP 6.6 μg mL^{-1} (0.05 U mL^{-1} at 60 °C towards 6C2FP). Thd conversion and 6C2FP-dR formation were calculated from [Thymine]/[Thd]$_{initial}$×100 % and [6C2FP-dR]/[R-1-P]$_{initial}$×100 %, respectively.

was rising steadily over 18 h and it rose faster at 60 °C than at 40 °C (Figure 3.8 A). For 6C2FP-dR synthesis, as expected, the product formation jumped to a maximum yield in a short time (69 % for 40 °C, 30 min; and 61 % for 60 °C, 5 min). Subsequently, the yield started to decline for the reaction at 60 °C, but for the reaction at 40 °C the yield maintained at 65 % even after 18 h (Figure 3.8 B). Interestingly, at 40 °C GtPNP became "alive" in the presence of dR-1-P & 6C2FP but was totally inactive for R-1-P & 6C2FP (Figure 3.6), which is obviously not due to the different enzyme concentrations, since the amount of GtPNP was 10 times higher in the case of R-1-P & 6C2FP than in the case of dR-1-P & 6C2FP.

It should be mentioned that the negative controls, with dR-1-P and 6C2FP but without enzyme, were performed simultaneously. No product (6C2FP-dR) was detected from them. Since our HPLC-assay cannot measure dR-1-P directly, an enzymatic method was figured out to check if dR-1-P was decomposed or hydrolysed in the controls. An incubated control (without enzyme) was analysed and no product was detected. Then GtPNP was added into in this control and incubated for 10 min at 40 or 60 °C. The reaction mixture was analysed again by HPLC. After adding GtPNP only 2 % 6C2FP-dR were formed (based on initial dR-1-P) in the control sample of 18 h incubation at 40 °C, thus *ca*. 3 % dR-1-P remained;

while no 6C2FP-dR was detected from the control sample run at 60 °C for 4 h, i.e. dR-1-P was totally decomposed after 4 h incubation at 60 °C.

In summary, for the synthesis of 2'-deoxynucleoside, low temperature and the time point of the maximum yield are the most important conditions to be considered. In the case of 6C2FP-dR, 40 °C appeared to be a suitable temperature for its synthesis because little decline of the product yield was observed from the second reaction. Thus, at 40 °C, a higher yield is expected in the one-pot transglycosylation and it can be regarded as an equilibrium-controlled reaction. If the reaction is performed at high temperature, an adjusted enzyme combination is needed, which allows dR-1-P to be transferred to the base as soon as possible, i.e. no excess dR-1-P in the system, then an improved product yield would be expected. It is advisable to study the degradation rate of dR-1-P in the reaction buffer at different temperatures and at different pHs, because these data can provide useful information for the determination of the process window.

3.3. Type3: Kinetically controlled reactions

Nucleosides with an F atom substitution at the 2'-position, namely 2'-dexoy-2'-fluoro *-ribo-* or *–arabino* nucleoside (FA, **16**; FanaA, **17**, respectively) are challenging compounds not only for the chemical synthesis but also for enzymatic synthesis. Both methods have been developed and a comprehensive overview and a careful comparison was published (Szeker 2012). A key point is that 2'-fluorinated pyrimidine nucleosides (including 2'-dexoy-2'-fluoro *–ribo* and *–arabino* nucleosides) can be synthesized efficiently by chemical methods (Liu et al. 2008; Howell et al. 1988; Turkman et al. 2010), while their purine counterparts are still suffering from tremendously complicated synthesis routes and separation steps (Thomas et al. 1994; Roshevskaia et al. 1986; Tennilä et al. 2000; Sivets et al. 2006). Luckily, a combination of a chemo-enzymatic approach can be designed to resolve these deficiencies. Explicitly, 2'-fluorinated pyrimidine nucleosides can be provided by chemical synthesis, which just offers the pentofuranosyl donor, for connecting it to a purine base by enzymatic transglycosylation, and then a desired purine nucleoside will be formed. A few examples of this chemo-enzymatic approach have been reported (Fernandez-Lucas et al. 2010; Yamada

et al. 2009; Zaitseva et al. 1999; Tuttle et al. 1993). A common limitation of the enzymatic approach is the low activity of the enzymes towards such modified substrates. Good news is that TtPyNP and all PNPs in the present study showed relative higher activity towards 2'-fluoronated nucleosides (within 10^{-2} U mg^{-1}, see Table 2.12) than the other reported enzymes such as TP and PNP from *E. coli* (Tuttle et al. 1993), PNP from *G. stearothermophilus* (Yamada et al. 2009) and NDT from *Lactobacillus reuteri* (Fernandez-Lucas et al. 2010). However, no superior product (FA and FanaA) yield over the other reported ones has been achieved by the enzymes in the present study, which might be attributed to the lower enzyme loading and shorter reaction time applied by us.

For the synthesis of FA and FanaA, the starting compounds are FU (**5**) and FanaU (**6**), respectively, with adenine as the purine base. In both reactions no equilibrium was observed, because of the limited catalysing ability of the enzymes, i.e. these are kinetically controlled reactions. Concretely, their transglycosylation process curves showed a steady and approximately linear increase of the product yield (synthesis) and a quicker rising curve of the uracil formation (phosphorolysis) (Szeker 2012). The difference of the two transglycosylation reactions is the substrate stability, which determines the reaction temperature and the fit to the temperature of the corresponding enzymes. It has been confirmed that above 55 °C it is evident that FU undergoes partly transformation into *O*^{2},2'-anhydro-1-(β-D-arabinofuranosyl)uracil (AnhU), and the latter will subsequently be hydrolysed and results in 1-(β-D-arabinofuranosyl)uracil (AraU, **7**) (paper I). Obviously, the reaction with FU should not be performed at a temperature above 55 °C; coincidently DgPNP has the highest activity towards 2'-fluoro nucleoside among the three PNPs and its optimal temperature is at 55 °C. Thus, TtPyNP and DgPNP are the best combination for FA synthesis (paper II). In contrast, FanaA is rather stable and higher temperature is beneficial for its phosphorolysis by TtPyNP. So TtPyNP and ApMTAP at 80 °C are a suitable combination for FanaA synthesis (paper II).

Thus, the synthesis of FA and FanaA were performed with 2 mM FU or FanaU and 1 mM adenine in 2 mM phosphate buffer (pH 6.5), enzyme loading was 0.1 mg mL^{-1} for each NPs, i.e. TtPyNP and DgPNP or ApMTAP. The product yields of FA and FanaA were very similar, i.e. a yield of ~24 % was achieved after 24 h

reaction at 55 °C or 80 °C based on the initial concentration of adenine (paper II).

For FA synthesis optimization was made by increasing the enzyme amount (TtPyNP 0.4 mg mL^{-1}; DgPNP 1 mg mL^{-1}) and extending the reaction time (up to 52 h). Moreover, the substrate ratio of FU to adenine was also considered. The process curves are shown in Figure 3.9. This time product formation is calculated based on the initial concentration of FU since adenine is in excess. On the whole, the substrate ratio has almost no influence on the product yield; while higher phosphate concentration (11 mM *vs* 8 mM) seems to have a slight positive effect, which resulted in a little higher product formation, notably higher uracil formation (higher FU conversion) and less undesired side product AnhU. These curves suggest that product formation was limited by the synthesis step, and not by the phosphorolysis step. The highest product yield was only 34 % after 52 h, which can be surely optimized further, for example by adding more DgPNP, changing the phosphate concentration, adjusting the pH, or by using immobilized enzymes.

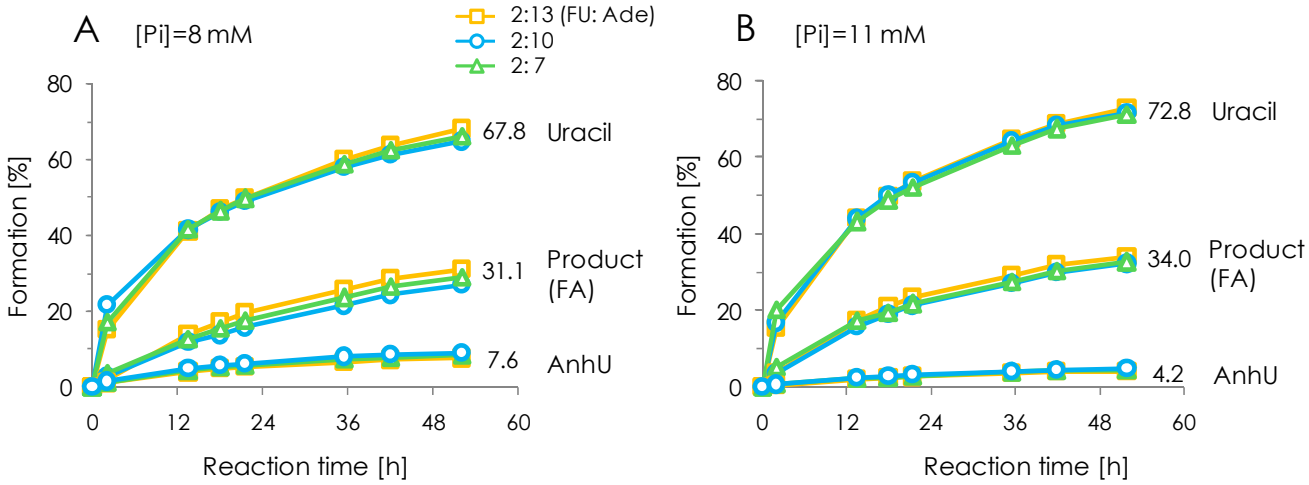


Figure 3.9. Progress curves of enzymatic synthesis of FA at 55 °C by TtPyNP (0.4 mg mL^{-1}) and DgPNP (1 mg mL^{-1}) with different substrates ratios (FU=2 mM, Ade=13, 10 and 7 mM). (A) 8 mM K-phosphate buffer, pH 7.0 (B) 11 mM K-phosphate buffer, pH 7.0. Formation of uracil, product (FA) and AnhU were calculated based on the initial concentration of FU. Formation values (%) from the reaction with substrate ratio of 2:13 (FU:Ade) were indicated beside the curves. AhU=O2,2'-anhydro-1-(β-D-arabinofuranosyl)uracil.

Chapter 4. Immobilization of thermostable NPs (paper III)

Nucleoside phosphorylases (NPs) are efficient biocatalysts, which have been extensively studied especially for R&D of modified nucleosides of pharmaceutical and biological interest (Mikhailopulo and Miroshnikov 2011; Mikhailopulo 2007). Thermostable NPs from thermophilic microorganisms, retaining the high *regio*- and *stereo*- selectivity of the mesophilic NPs, have further advantages that are not only associated with an increased thermostability but also with a broader substrate spectrum and a wider tolerance to many stresses (Zhou et al. 2013; Szeker et al. 2012; Almendros et al. 2012; Taran et al. 2009; Utagawa et al. 1985). Among them, purine nucleoside phosphorylase (EC 2.4.2.1) from *Geobacillus thermoglucosidasius* (GtPNP) (Zhou et al. 2013) and pyrimidine nucleoside phosphorylase (EC 2.4.2.2) from *Thermus thermophilus* (TtPyNP) (Szeker et al. 2012) have shown such promising properties. However, the majority of modified nucleosides are not currently synthesized by biocatalysts yet, but rather by traditional chemical synthesis, which consists of laborious multistage processes and often leads to low yields due to the poor *regio*- and *stereo*- selectivity resulting in the formation of unwanted isomers. It is conceivable that with the development of biocatalysts, enzymatic synthesis could be adopted for the biocatalytic synthesis of valuable pharmaceutical intermediates for their economic industrial production.

Immobilization has been recognized as a very powerful tool to realize the application of enzymes under the relatively harsh conditions required in some industrial processes. More importantly, immobilized enzymes can be easily recovered for a potential reuse, which permits the reduction of biocatalyst cost, simplifies the design of the reactor and allows to realize in some cases a better control of the reaction (Illanes 2008; Mateo et al. 2007b). Therefore, it is highly desirable to immobilize NPs even if they are thermostable.

There are several reports on the immobilization of NPs on different supports. For instance, thermostable GsPNP and GsPyNP from *Geobacillus stearothermophilus* (previously *Bacillus stearothermophilus*) were immobilized on anion exchange resins (DEAE-Toyopearl 650 M) (Hori et al. 1991b), or on the aminopropylated

macroporous glass AP-CPG-170 (Taran et al. 2009); UP and PNP from *E. coli* were immobilized on epoxy-activated resin Sepabead EC-EP (Zuffi et al. 2004); EcTP from *E. coli*, BsPNP and BsPyNP from *Bacillus subtilis*, AhPNP from *Aeromonas hydrophila et al.*, were immobilized and compared with different ionic supports or epoxy-activated resins, with or without cross-linking agents (Serra et al. 2013b; Serra et al. 2013a; Serra et al. 2011; Rocchietti et al. 2004). To sum up, the epoxy supports coated with polyethyleneimine (PEI) resulted in a more stable biocatalyst with a moderate residual activity. However, several common drawbacks were identified including: (i) the potential dissociation of the multimeric NPs leads to a loss of enzyme activity and contamination of the product, which is a critical problem in pharmaceutical processes (Fernandez-Lafuente 2009); (ii) the final biocatalyst has a low volumetric activity because of the relatively low enzyme loading capacity of the supports (Brady and Jordaan 2009).

MagReSyn™ Epoxide Microspheres (MagEM), emulsion-derived particles, are an ingenious combination of permanent magnetic magnetite particles and a hydrophilic polymer with epoxide-functional groups and ionic character (Jordaan et al. 2013). This invention provides particles with adequate surface area and high functional group density throughout the open fibre network. It has been shown that the porous particles allow for stabilization of multimeric enzymes through multipoint attachment of the subunits, preventing subunit dissociation (Twala et al. 2012). This makes it a potentially ideal support for the immobilization of multimeric enzymes (Fernandez-Lafuente 2009). The combination of ionic and covalent binding force are coincidentally considered to be the appropriate supports for the immobilization of NPs (Serra et al. 2013b; Serra et al. 2013a; Serra et al. 2011; Rocchietti et al. 2004).

The above arguments prompted us to explore the immobilized enzymes and in this paper we report the efficient immobilization of the multimeric TtPyNP and GtPNP on MagEM, and the characterisation of the respective preparations. Finally, the optimal co-immobilized biocatalysts were used for the synthesis of 2,6-dihalogenated purine nucleosides, and the production yields were optimized.

4.1. Experimental

4.1.1. General

All chemicals and solvents were of analytical grade or higher and purchased, if not stated otherwise, from Sigma-Aldrich (Steinheim, Germany), Carl Roth (Karlsruhe, Germany), TCI Deutschland (Eschborn, Germany), and VWR (Darmstadt, Germany). Water was purified by a purification system from Merck Millipore (Schwalbach, Germany). MagEM and the magnetic separator were kindly donated by ReSyn Biosciences (Pretoria, South Africa).

PyNP from *Thermus thermophilus* (TtPyNP), and GtPNP from *Geobacillus thermoglucosidasius* (GtPNP) were expressed in *E. coli* BL21 and purified by heat-treatment and standard Ni-NTA affinity chromatography as previously described (Szeker et al. 2012; Zhou et al. 2013). The concentrations of purified protein used for immobilization were determined by NanoDrop1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, USA) (Desjardins et al. 2001), which were 0.53 and 1.58 mg mL^{-1} of TtPyNP, and 6.61 mg $mL^{-1}$of GtPNP. All the enzymatic reactions were performed on a thermomixer (Biozym, Hessisch Oldendorf, Germany) and analysed by HPLC Agilent 1200 series system equipped with an Agilent DAD detector. ARP C18 column of Phenomenex Gemini 5 µm, 150 × 4.60 mm was used for chromatography (Torrance, United States). Immobilized enzymes were incubated in PCR machine (Eppendorf, Hamburg, Germany) for the thermostability test.

4.1.2. Procedures of enzyme immobilization

TtPyNP and GtPNP were immobilized on MagEM separately according to the manual (http://www.resynbio.com) with a slight modification. Briefly, 20 µL suspension beads (25 mg mL^{-1}) were washed and equilibrated in binding buffer (50 mM K-phosphate buffer, pH 6) shortly before using. Enzyme solution with the binding buffer was adjusted to 10 times volume of the initial beads suspension, i.e. 200 µL of total volume. The enzyme and beads suspension were mixed thoroughly and were put on the thermomixer at 50 °C and shaking at 950 rpm for 4 h. After binding procedure, the beads were collected by magnetic separator and the supernatant was discarded or used for quantification of immobilization

yield. The beads were washed with washing buffer (50 mM K-phosphate buffer, pH 7.0; 1 M NaCl) for three times to remove non-covalently bound enzyme. Residual epoxide groups were quenched using a similar procedure, except that a quenching solution (1 M Tris pH 9.0) instead of the binding solution was used with incubation for 1.5 h. The beads were washed with 50 mM K-phosphate buffer pH 7.0 for three times and were ready to use or were stored at 4 °C.

After the optimal conditions were elaborated, co-immobilization of TtPyNP and GtPNP was performed under these conditions in the same way of the individual enzyme immobilization, but the coupling enzymes were first admixed and then added into the beads in the binding step.

4.1.3. Enzyme activity assay

The activities of free enzyme and the supernatant from the binding step (soluble enzyme) were determined as previously described (Szeker et al. 2012; Zhou et al. 2013). Briefly, in a 200 µl reaction mixture, 1 mM substrate was dissolved in 50 mM potassium phosphate buffer (pH 7.0) and pre-heated at a defined temperature on a thermomixer (at 300 rpm), a suitable amount of enzyme was added so that the reaction rate was linear over time. The reaction was stopped by adding 0.5 volume of 10 % trichloroacetic acid. After centrifugation, the supernatant was analysed by HPLC.

Activity measurements with the immobilized enzymes were performed similarly, but at a high shaking speed (1000 rpm) with a reaction volume of 200 or 500 µl on a thermomixer. Reaction was stopped by using a permanent magnet stand to separate the beads and the solution, the latter was withdrawn for analysis.

For soluble and immobilized TtPyNP, 0.8 mM uridine (Urd) at 80 °C and 2 mM Urd at 40 °C were applied in the activity test. For soluble GtPNP, 1 mM inosine (Ino) at 70 °C was used; while for immobilized GtPNP, 1 mM cytidine (Cyd) at 70 °C, 2'-deoxycytidine (dCyd) at 70 °C, 2'-amino-2'-deoxyinosine (aIno) at 30–60 °C, and 2'-amino-2'-deoxyadenosine (aAdo) at 30–80 °C were used in activity tests. For co-immobilized TtPyNP and GtPNP, 5 mM Urd and 3.5 mM 2,6-dichloropurine (26DCP) in 2 mM K-phosphate buffer were used as the substrates.

All assays were performed at least in duplicates. HPLC analysis was carried out as previously described (Szeker et al. 2012). The conversions were quantified as percentage of converted substrate from the HPLC results. One unit (U) of enzyme activity was defined as the quantity of the enzyme required to catalyse 1 µmol substrate per minute under the reaction conditions.

4.1.4. Immobilization yield (expressed activity)

Immobilization yield was defined as the percentage activity of the immobilized enzyme to the native enzyme, also referred to the residual activity.

4.1.5. Screening of immobilization conditions

Immobilization was conducted following the procedure described above, if not otherwise stated.

Screening for binding buffer pH

Binding buffers with pH 5.0–8.5 (50 mM K-phosphate) or pH 8.5–9.0 (50 mM triethanolamine) were applied, wherein 70 µL TtPyNP stock (0.53 mg mL^{-1}) and 20 µL GtPNP stock (6.61 mg mL^{-1}) were loaded.

Screening for enzyme loading

Two concentrations of TtPyNP were applied, with 1.8–14.2 mg enzyme per mL suspension beads (i.e. "mg mL^{-1} beads" for short, same as below), which equals to 0.07–0.57 g per g dry beads; while GtPNP was applied in a range of 2–40 mg mL^{-1} beads, which equals to 0.09–1.59 g per g dry beads. As binding buffer 50 mM K-phosphate (pH 6.5) was used here.

Design of Experiments (DoE)

To explore the influential effects in the immobilization procedure, a two-level-three-factor full factorial design was employed. The three factors and their levels were: binding time (1–4 h), binding temperature (25–50 °C), and quenching time (0–2 h). Immobilization yield (%) was the response, which needed to be maximized. The experimental matrix is shown in Table 4.1.

GtPNP was used in this DoE as a representative. The constant conditions were: binding buffer 50 mM K-phosphate (pH 6.0), enzyme loading 7 mg mL^{-1} beads, binding buffer volume 100 µL (with 10 µL suspension beads).

Experiments design and data analysis were done by the software MODDE 10.0.0 (Umetrics AB, Malmö, Sweden).

4.1.6. Temperature profile of immobilized GtPNP

Temperature optima of free and immobilized GtPNP (6.6 mg mL^{-1}beads) were determined from the activities at temperatures ranging from 30 to 80 °C towards 1 mM aAdo or aIno (in 50 mM K-phosphate buffer pH 7.0). Activities were assayed as stated above.

4.1.7. Thermostability of immobilized enzymes

Immobilized TtPyNP (14.2 mg mL^{-1} beads), GtPNP (23.1 mg mL^{-1} beads) as well as co-immobilized TtPyNP and GtPNP (4.7 and 16.5 mg mL^{-1} beads, respectively) were incubated with 50 mM K-phosphate buffer (1 mg beads per mL buffer) in a PCR machine for 20 h at 75 or 80 °C. The residual activity was determined at different incubation time following the activity assay described above.

4.1.8. Synthesis of 2,6-dihalogenated purine nucleosides

2,6-Dichloro-9-(β-D-ribofuranosyl)purine (26DCP-R)

Two-pot reaction: a reaction mixture (200 µL) containing 0.69 mg immobilized TtPyNP-beads (0.19 mg TtPyNP, *ca.* 4.1 U), 20 or 40 mM Urd, and 2 or 10mM K-phosphate buffer (pH 7.0), was stirred at 1000 rpm at 80 °C. When the equilibrium was reached (i.e. where Urd conversion became constant), the reaction mixture and beads were separated and the supernatant was added into the same volume mixture (200 µL) containing 1.16 mg immobilized GtPNP-beads (0.66 mg GtPNP, *ca.* 2.4 U), 15 mM 26DCP, and 2 or 10 mM K-phosphate buffer (pH 7.0). The reaction was performed at 70 °C stirring at 1000 rpm.

One-pot reaction: a reaction mixture (400 µL) containing 1.85 mg co-immobilized TtPyNP and GtPNP (0.19 mg TtPyNP, *ca.* 4.1 U; 0.66 mg GtPNP, *ca.* 2.4 U), 10 or 20

mM Urd, 7.5 mM 26DCP, and 2 or 10 mM K-phosphate buffer (pH 7.0), was stirred at 1000 rpm at 70 °C.

Samples were withdrawn and diluted 20 times after removing the magnetic beads. The reaction progress was monitored by HPLC.

One unit (U) of TtPyNP was defined as the amount of immobilized TtPyNP required to convert1 µmol Urd to uracil in 50 mM K-phosphate (pH 7.0) per minuteat 70 °C; one Uof GtPNP was defined as the amount of immobilized GtPNP converts 1 µmol 26DCP to 26DCP-R in 2 mM K-phosphate (pH 7.0) per minute at 70 °C.

6-Chloro-2-fluoro-9-(β-D-ribofuranosyl)purine (6C2FP-R)

A reaction mixture (400 µL) containing: 1.85 mg co-immobilized TtPyNP and GtPNP (0.19 mg TtPyNP, *ca.*4.1 U; 0.66 mg GtPNP, *ca.* 5.4 U), 20-50 mM Urd, 20mM 6C2FP, and 2-10 mM K-phosphate buffer (pH 7.0), was stirred at 1000 rpm at 70 °C. Samples were taken as described above.

The unit of TtPyNP was defined as above; one U GtPNP was defined as the amount of immobilized GtPNP required to convert 1 µmol 6C2FP to 6C2FP-R in 2 mM K-phosphate (pH 7.0) per minute at 70 °C.

4.2. Results

4.2.1. Screening of immobilization conditions

Influence of binding buffer pH and the "maturing" of GtPNP

For the epoxide support, although high pH was expected to favour the immobilization capacity (Mateo et al. 2007a; Mateo et al. 2002), TtPyNP displayed higher residual activity in binding buffer with lower pH values, and the related supernatant activities agreed with the residual activities (Figure 4.1A). Surprisingly, with binding buffer pH 5-6 the TtPyNP immobilization yields were above 150 %, indicative of enzyme activation.

Unlike TtPyNP, GtPNP did not show a clear preference for the binding buffer pH (Figure 4.1C). The immobilization yields of GtPNP were tested with two substrates: aIno and dCyd, which yielded similar results of 30-60 % conversion (Figure 4.1B and C; black bar) for all binding buffers tested. The corresponding supernatants,

however, showed the residual activities were all less than 1.5 % (data not shown). Since the total residual activities from the immobilized beads (i.e. immobilization yield) and from the supernatant should be 100 %.

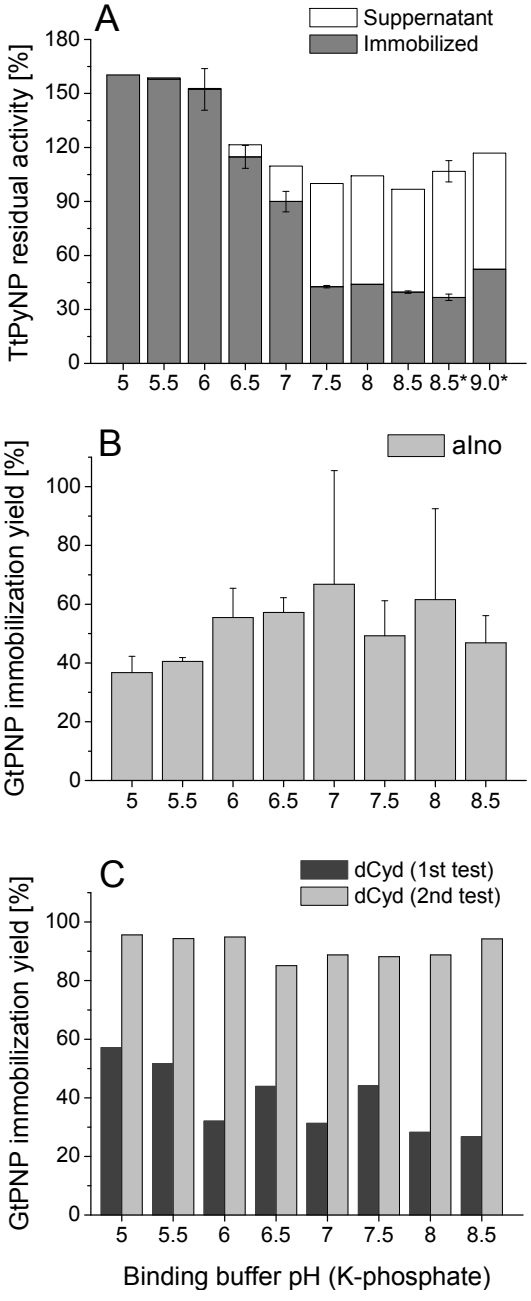


*Figure 4.1. Influence of pH of binding buffer (50 mM K-phosphate; *50 mM triethanolamine) on the residual activity of immobilized TtPyNP (A, gray bar: 2 mM Urd at 40 °C) and the supernatants (A, white bar: 0.8 mM Urd at 80 °C), immobilized GtPNP (B: 1 mM aIno at 30 °C; C: 1 mM dCyd at 70 °C). TtPyNP loading: 1.8 mg mL^{-1} beads; GtPNP: 7 mg mL^{-1} beads.*

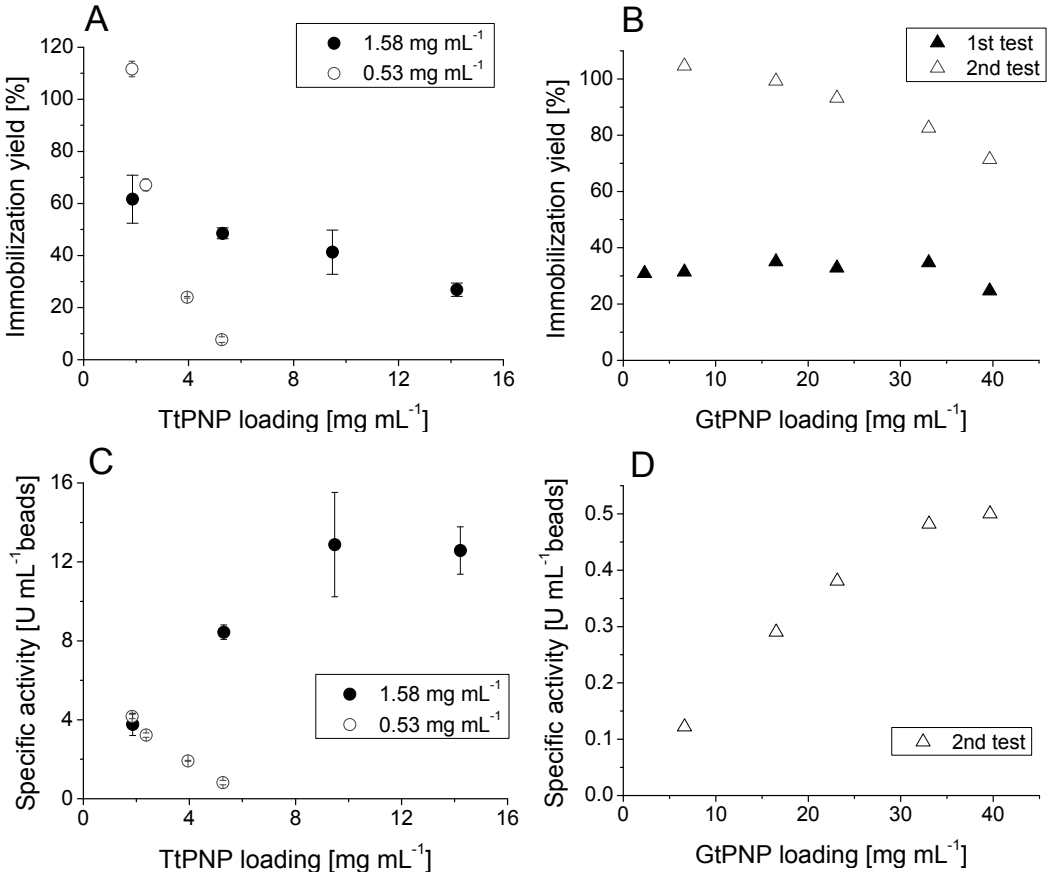


Figure 4.2. Influence of enzyme loading on the immobilization yield (A: TtPyNP; B: GtPNP) and the volumetric specific activity (C: TtPyNP; D: GtPNP). 1.58 mg mL-1 and 0.53 mg L-1 TtPyNP stocks were applied (A, C). The 2nd test of GtPNP samples were run after 28 days of the 1st test (B, D). Binding buffer: 50 mM K-phosphate (pH 6.5). TtPyNP activities were tested towards 2 mM Urd at 40 °C; GtPNP activities were tested towards 1 mM dCyd at 70 °C.

Thus, it seems that most GtPNP was immobilized, but with activity loss. Unexpectedly, the second run of the same GtPNP samples tested after 13 days storage showed dramatic increasing of their activities (Figure 4.1 C, gray bars), which reached 88-96 % yields in agreement with the supernatant data. In addition, a similar phenomenon was observed when different amounts of GtPNP were loaded (Figure 4.2 B). However, a similar behaviour was not observed for the TtPyNP samples, which kept the same activity in the subsequent second and the third reactions.

Influence of enzyme loading

As more enzymes are loaded onto the immobilization support, it would be expected that the immobilization yield (residual activity) would decrease due to mass transfer limitations. And with more enzyme loading the biocatalyst specific activity will increase at beginning and stop or decrease when the supports get saturated with enzyme. In this study, with higher loading the immobilization yield of TtPyNP decreased while the specific activity kept increasing when the high concentration of the enzyme stock solution (1.58 mg mL^{-1}) was applied up to 9.5 mg mL^{-1} beads (Figure 4.2 A, C; closed circles). By contrast, when a lower concentration of TtPyNP (0.53 mg mL^{-1}) was applied, both the immobilization yield and the specific activity decreased dramatically with the enzyme loading (Figure 4.2 A, C; open circles). With the same enzyme loading (e.g. TtPyNP was loaded as 5.3 mg mL^{-1}, see Figure 4.2 A, C), a higher stock concentration resulted in a higher immobilization yield as well as in a higher specific activity. Therefore, to obtain immobilized enzyme with high yield and activity, a high concentration of the enzyme stock solution was required. GtPNP was loaded up to 33 mg per mL suspension beads and its immobilization yield of the 2^{nd} test was still above 80 % (Figure 4.2 B; open triangles). Saturation was observed when GtPNP was loaded at 40 mg mL^{-1}; the specific activity stopped to increase at this point (Figure 4.2 D; open triangles). In summary, TtPyNP can be loaded up to 9.5 mg mL^{-1} (equals to 0.38 g per g dry beads) to obtain an optimal specific activity and an immobilization yield of 41 %; GtPNP can be applied up to 33 mg mL^{-1} (equals to 1.32 g per g dry beads) with an immobilization yield of 83 %.

Design of Experiments (DoE)

The influence of three factors (binding time, binding temperature, and quenching time) on the GtPNP immobilization yield was investigated by DoE using a two-level three-factor (2^3) full factorial design. The worksheet and the results are shown in Table 4.1.

Binding temperature and binding time show both strong positive effects (Figure 4.3), which are related to the reaction yield of the protein residues that bind to the epoxide groups of the beads. In contrast, the effect of quenching was insignificant and thus was removed from the model. Hence, a quadratic model

was established using the multiple linear regression method by software MODDE 10.0.0. The results suggest that higher temperature and longer binding time within the tested range have positive effects on the immobilization yield. The dependencies of immobilization yield on binding temperature and binding time according to the model are summarized in the contour plots (Figure 4.4). So far, the highest immobilization yield of 65.7 % was obtained with a binding temperature at 50 °C and a binding time of 4 h. However, with higher binding temperature and longer binding time, the immobilization yield could be reached higher.

Table 4.1. Immobilization conditions for GtPNP with immobilization yield as the response.

Exp. No.	Binding time [h]	Binding temp. [°C]	Quenching time [h]	Immobilization yield [%][a]
1	1	25	0	17.1
2	4	25	0	38.3
3	1	50	0	59.7
4	4	50	0	64.5
5	1	25	2	16.6
6	4	25	2	41.4
7	1	50	2	57.1
8	4	50	2	65.7
9	2.5	37.5	1	50.8
10	2.5	37.5	1	53.1
11	2.5	37.5	1	52.8

[a] Immobilization yield was measured after one day "maturing" using 1 mM dCyd, 50 mM K-phosphate as substrates and 70 °C as the reaction temperature.

The parameters used for the model evaluation showed that the model (i) fitted the measurements well (R^2=0.996), (ii) could predict new data well (Q^2=0.985), (iii) was valid because the model error was in the same range as the replicate error (model validity=0.764), and (iv) showed a good reproducibility (=0.995). For more details, see supplementary information (Fig. S1, Table S1).

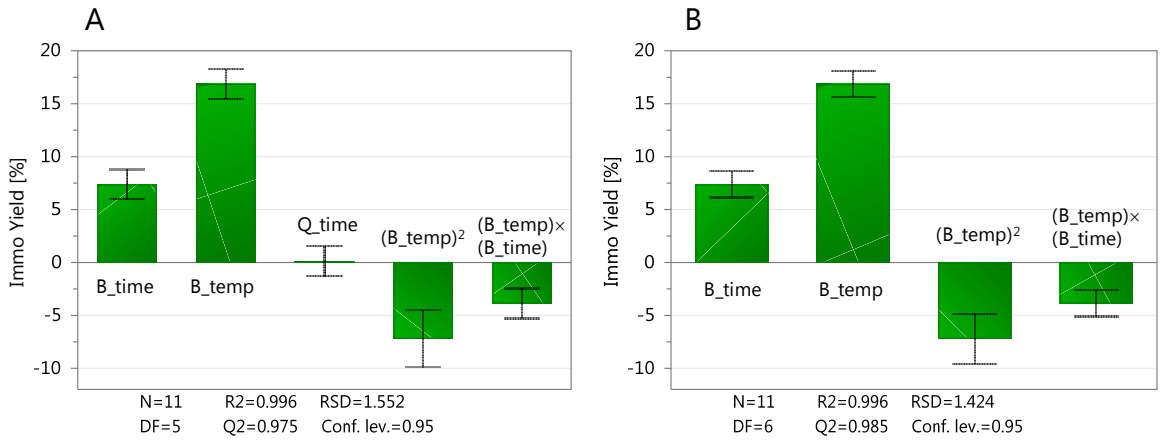


Figure 4.3. Normalized coefficient plot of Immobilization Yield regression model with (A) and without (B) quenching time as a factor. Binding time (B_time), binding temperature (B_temp), and quenching time (Q_time).

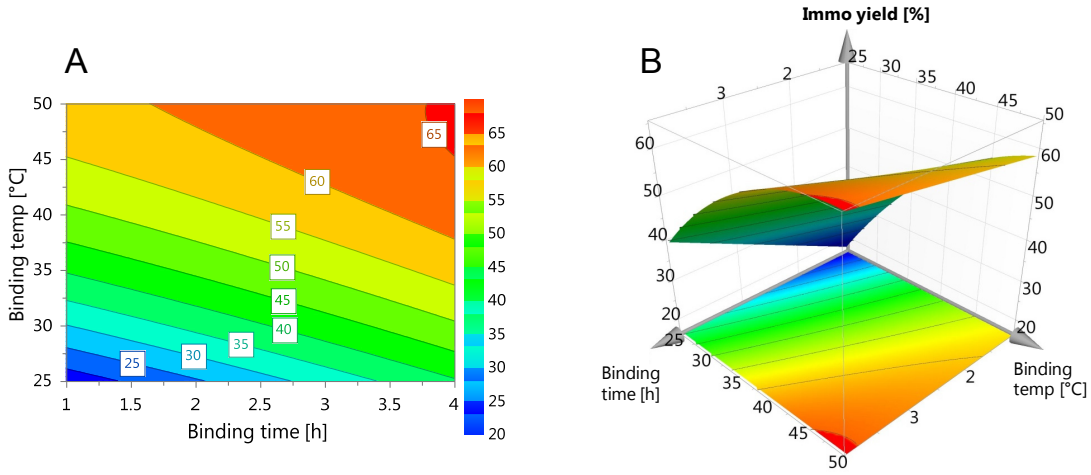


Figure 4.4. Immobilization yield dependency on binding temperature and binding time. A: 2D contour plot; B: 3D surface plot.

4.2.2. Temperature profile of the immobilized GtPNP

To investigate the influence of immobilization on the enzymatic activity, the temperature dependence of immobilized GtPNP activity was studied for two substrates (aIno and aAdo). The temperature profiles for the free and immobilized GtPNP towards aIno are illustrated in Figure 4.5. When the substrate was aAdo, very similar profiles were observed (data now shown). As displayed in Figure 4.5, both free and immobilized GtPNP have the optimal temperature at 70 °C as it was observed before with Ino as substrate (Zhou et al. 2013).

Interestingly, the relative activity, i.e. the expressed activity, of immobilized GtPNP was increasing with the temperature from 68 % at 30 °C to 111 % at 80 °C. These data imply that the immobilization would not hinder the enzyme's thermo-activity, but favour the enzyme activity especially at high temperatures.

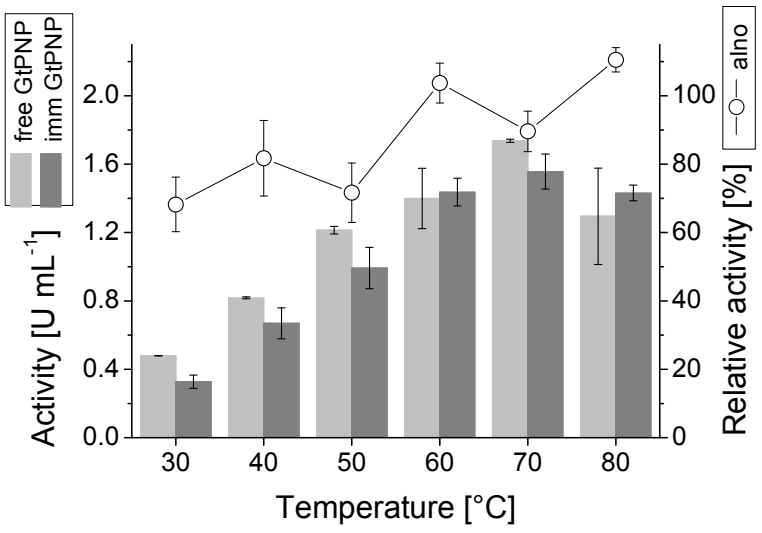


Figure 4.5. Temperature dependence of the specific activity of the free and immobilized GtPNP towards alno and the relative activity of immobilized GtPNP vs free GtPNP.

4.2.3. Thermostability of immobilized enzymes

The stability of immobilized TtPyNP, GtPNP and co-immobilized TtPyNP & GtPNP were tested at high temperature (80 °C and 75 °C) for an incubation over 20 h. The immobilized TtPyNP showed a high residual activity of 96 % after 20 h at 80 °C (Figure 4.6), which was higher than its native counterpart (75 % residual activity, 23 h, 80 °C (Szeker et al. 2012)). Moreover, immobilized GtPNP was incubated at 75 °C, which is 5 °C higher than its optimal temperature, and the calculated half-life was 24.1 h (75 °C) (Figure 4.6), while the native GtPNP at 75 °C has a half-life of only 6.3 h. The stability of co-immobilized TtPyNP and GtPNP was very similar to the immobilized GtPNP; both had a residual activity of 58 % after 20 h incubation at 75 °C.

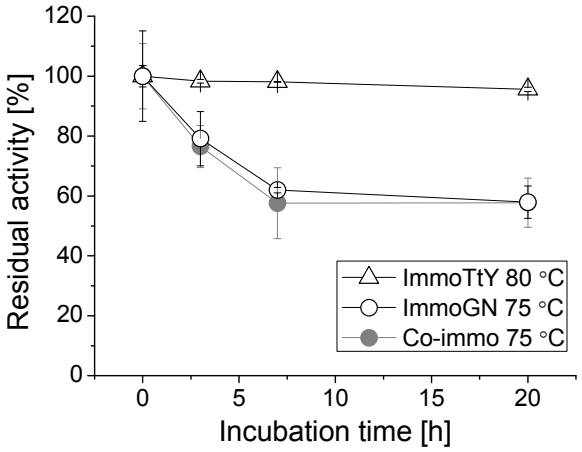
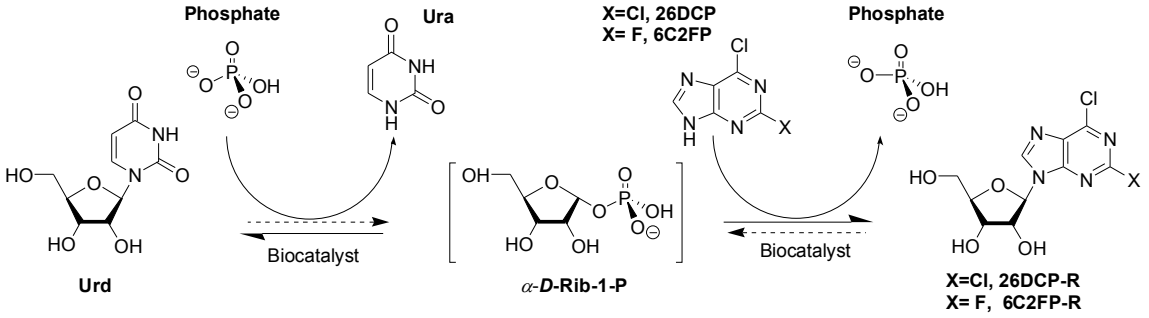


Figure 4.6. Thermostability of immobilized TtPyNP at 80 °C (open triangle), immobilized GtPNP at 75 °C (open circle), and co-immobilized TtPyNP and GtPNP at 75 °C (close circle).

4.2.4. Synthesis of 2,6-dihalogenated purine nucleosides

The optimal immobilized enzymes were applied under different reaction conditions to optimize the synthesis of 2,6-dihalogenated purine nucleosides. The basic synthetic strategy is illustrated in Scheme 4.1.



Scheme 4.1. Synthesis of 2,6-dihalogenated purine nucleosides by biocatalyst.

The use of magnetic beads allowed for the simple separation of biocatalysts from the reaction mixture making it possible to compare two synthesis strategies: (i) Two-pot reaction. Firstly, α-*D*-Rib-1P was produced from Urd by immo_TtPyNP; secondly, the reaction mixture containing α-*D*-Rib-1P was separated from immo_TtPyNP and used as the glycosyl donor for the product (26DCP-R) synthesis with added immo_GtPNP and purine base (26DCP). (ii) One-pot reaction. All substrates and biocatalysts were added into a single reaction mixture.

Table 4.2. Application of biocatalyst for the synthesis of 26DCP-R[a].

Biocatalyst	Vial	Substrate [mM]		Conversion [%][b]	
		Urd	Phosphate	Urd	26DCP
Immo_TtPyNP[c]	1	20	2	14.7	-
	2	20	10	23.7	-
	3	40	10	15.5	-
Immo_GtPNP[d]	4	10	2	14.8	17.0
	5	10	10	23.6	26.7
	6	20	10	15.5	36.5
Co-immo[e] (TtPyNP&GtPNP)	7	10	10	57.0	55.6
	8	10	6	54.4	60.3
	9	10	2	53.0	63.2
	10	10	1.4	57.1	68.7
	11	20	1.4	36.1	78.5
	12	20	6	37.6	72.8

[a] Vials 1-12 were 30 min reactions (equilibrium reached). Experimental conditions were described under Experimental section 4.1.8. The highest standard deviation is less than 10 %.

[b] Conversion was calculated from HPLC analysis, which were the percentage of the converted Urd or 26DCP.

[c] First step of "two-pot reaction" for producing α-D-Rib-1P for the second step. Vials 1-3 were the phosphorolysis of Urd (without 26DCP).

[d] Second step of "two-pot reaction" for 26DCP-R synthesis. Vials 4-6 were from the reaction mixture of vials 1-3 added with immo_GtPNP and 26DCP (final 7.5 mM).

[e] One-pot reaction. Vials 7-12 contain 7.5 mM 26DCP.

As reported in Table 4.2, for the two-pot reaction, on increasing phosphate concentration the yield (conversion of purine base by HPLC analysis, the same as below) increased (17.0 % to 26.7 %, vials 4 and 5). In contrast, for the one-pot reaction the yield increased with decreasing phosphate concentration (55.6 % to 68.7 %, vials 7-10; and 72.8 % to 78.5 %, vials 11-12). Moreover, with the same substrates composition, the one-pot reaction resulted in a significantly improved yield compared to the two-pot reaction (63.2 % *vs* 17.0 %, vial 9 *vs* 4; 55 % *vs* 26.7 %, vial 7 *vs* 5). A yield improvement was also achieved, as expected, by raising the concentration of Urd (vial 6 *vs* 5 and 11 *vs* 10). The highest yield of 26DCP-R (78.5 %) was obtained from the one-pot reaction by co-immobilized TtPyNP and GtPNP (10.3 U mL^{-1} and 5.9 U mL^{-1}) at 70 °C with 1.4 mM phosphate, 20 mM Urd and 7.5 mM 26DCP in 30 min. The productivity of 26DCP-R was calculated as 1.51 g L^{-1} h^{-1}.

According to the results of 26DCP-R synthesis, one-pot reaction with low phosphate concentration and high substrate concentration were used for 6C2FP-R synthesis because 6C2FP has higher solubility than 26DCP. For the synthesis of 6C2FP-R (Figure 4.7), the highest yield (85.5 %) was obtained from one-pot reaction by co-immobilized TtPyNP and GtPNP (10.3 U mL^{-1} and 13.5 U mL^{-1}) at 70 °C with 2 mM phosphate, 50 mM Urd and 20mM 6C2FP in 180 min. The productivity of 6C2FP-R was calculated as 1.97 g L^{-1} h^{-1}.

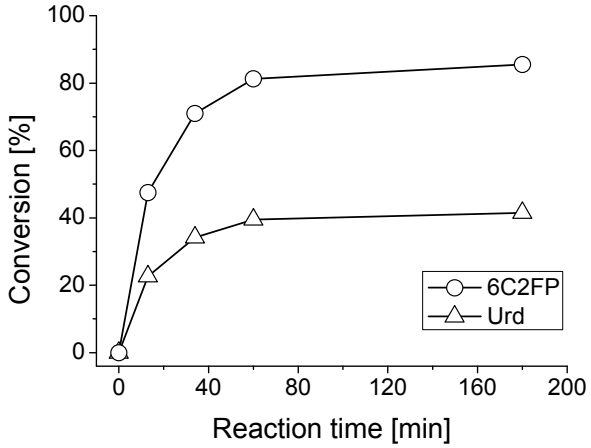


Figure 4.7. The course of synthesis of 6C2FP-R by co-immobilized TtPyNP and GtPNP)10.3 UmL^{-1} and 13.5 U mL^{-1}) at 70 °C in 2 mM phosphate buffer (pH 7.0) with 50 mM Urd and 20 mM 6C2FP.

4.3. Discussion

In this study, the main purpose of enzyme immobilization is to improve the efficiency of enzyme utilization. Therefore, high immobilization yield with high enzyme loading is the goal. Considering the immobilization protocol and the beads character, binding buffer pH, enzyme loading, binding time and temperature, as well as quenching were identified as the main factors to be investigated. Among these, the first two have dominant influence. The pH of the binding buffer will affect the reactivity of the enzymes amino acid residues to the epoxide functional groups on the beads; while the enzyme loading affects the mass or volumetric specific activity of the immobilized enzymes (Illanes 2008). Thus, a one-shot design was employed to determine the influence of the binding

buffer pH and enzyme loading and factorial design was applied for determining the influence of the other factors.

The functional groups of the MagReSyn™ Epoxide Microspheres (MagEM) are epoxides, which are able to react with the nucleophile groups on the enzyme surface (e.g. amino, thiol, phenolic and imidazole) under very mild experimental conditions (e.g. pH 7.0) for a long period of time (Mateo et al. 2007a; Katchalski-Katzir and Kraemer 2000). At higher temperatures and at alkaline pH conditions, the reactivity of the enzyme residues is supposed to be higher as most of the nucleophile groups on the enzyme surface have a pKa above 7.0 (Mateo et al. 2007a; Mateo et al. 2002). To shorten the incubation time, we chose 50 °C as the standard immobilization temperature since both TtPyNP and GtPNP are thermostable enzymes and are stable at 80 °C and 70 °C, respectively (Zhou et al. 2013; Szeker et al. 2012).

The conventional epoxy supports are mostly hydrophobic and require high ionic strength (e.g. 1-3 M sodium phosphate) to improve the hydrophobic interaction between protein and the support. It was reported that at low ionic strength (50 mM sodium phosphate) the enzyme penicillin G acylase couldn't be immobilized to any of the commercial epoxy supports Eupergit C, Sepabeads-EP or agarose epoxy (Mateo et al. 2002). Moreover, using Eupergic C or Sepabeads FP-EC3, *Bacillus subtilis* UP and PNP could be absorbed on the supports at high ionic strength (2 M potassium phosphate buffer), however, the expressed activity was very poor to negligible (Rocchietti et al. 2004). The inactivity could be ascribed to the exposure of the enzyme active center onto the hydrophobic support by the force of adsorption in the high ionic strength medium (Mateo et al. 2002). By contrast, the immobilization of the above mentioned BsUP and BsPNP using modified Sepabeads FP-EC3 with PEI coated by dextran aldehyde and hydrophilic glyoxyl-agarose, the expressed activity increased to 78 % and 50%, respectively (Rocchietti et al. 2004). Therefore it is advisable to use a hydrophilic epoxide support for the immobilization of NPs; otherwise the distortion of the enzyme active center might occur by using hydrophobic support in high ionic strength buffer. The polymer matrix chosen for this study (epoxy activated hydrophilic polymer) allowed for the use of low ionic strength buffers (50 mM phosphate) and provided the highest chance for immobilization success.

Immobilized GtPNP was found to increase in activity during storage within 15-30 days. This unique phenomenon has not been described previously in the literature. Some possible explanations could be: (i) The immobilized enzyme needs time for its "maturing" or rearrangement. (ii) The cationic nature of the support may adsorb phosphate anions from the storing solution, which can increase the apparent activity of the enzyme. This phenomenon was, however, not observed in the TtPyNP samples. Besides, the preference of the binding buffer pH was also different for GtPNP and TtPyNP (Figure 4.1).

The divergence can potentially be ascribed to the structural differences of GtPNP and TtPyNP. According to Pugmire and Ealick (Pugmire and Ealick 2002), GtPNP belongs to the PNPII subfamily of the NP-I family, which is hexameric with single-domain subunits; TtPyNP belongs to the NP-II family, which is dimeric with two-domain subunits. Furthermore, TtPyNP requires significant domain movements for catalysis, while GtPNP does not. It appears that TtPyNP needs a flexible support for the immobilization to allow its domain movements and maintenance of its activity. The MagEM used in this study can fulfil this requirement through the claimed flexible polymer structure of the support. Indeed, a similar strategy was employed successfully for the immobilization of *E. coli* TP (Serra et al. 2011) and *B. Subtilis* PyNP (previously BsUP) (Serra et al. 2013a; Rocchietti et al. 2004), which belong to the same enzyme family, onto epoxide Sepabeads which were coated with PEI and finally cross-linked with dextran. In contrast, GtPNP is relatively rigid and if it was immobilized on flexible support it might need time to recover its rigid structure. PNP from other sources, e.g. *Aeromonas hydrophila* PNPII (Serra et al. 2013b) and BsPNP (Rocchietti et al. 2004), was immobilized by different strategies. The best result, a high immobilization yield of 20-30 % and a high stability (residual activity >60 % after 20 h), was obtained by the use of a hydrophilic support, glyoxyl-agarose at pH 10, and addition of a surface-active agent (e.g. Triton X-100 or glycerol) and an enzyme ligand (Serra et al. 2013b; Rocchietti et al. 2004). By comparison, immobilization on MagEM at mild conditions (50 mM phosphate buffer, pH 6.5) without any additive or post-immobilization quenching, the immobilization yield of GtPNP was above 80 % (Figure 4.2 B, 2nd test) even with very high enzyme loading of 1.32 g enzyme per gram of dry beads; while the highest previously

reported loading of AhPNPII was only 5 mg g^{-1} (Serra et al. 2013b). Furthermore, the immobilized GtPNP maintained 58 % activity after 20 h at 75 °C (Figure 4.6), which indicates the suitability of the system for operation at higher temperature to increase reaction speed.

The multimeric TtPyNP and GtPNP were efficiently loaded up to 0.38 and 1.32 g per g dry beads with 41 % and 83 % immobilization yield, respectively. This unusually high binding capacity was also reported on aldehyde ReSyn™ microspheres for other proteins such as BSA (1.7 g per g beads), NOD (1.3 g per g beads, 24 % immobilization yield) as well as for the multimeric enzyme GDH (0.5 g per g beads, 42 % immobilization yield) (Twala et al. 2012).

By contrast, the NPs family enzymes EcUP and EcPNP immobilized on amino propylated macroporous glass AP-CPG-170 was loaded at 3 mg g^{-1} (Konstantinova et al. 2004); BsPyNP immobilized on Sepabeads-PEI-dextran and BsPNPon glyoxyl-agarose (Rocchietti et al. 2004), CpUP, CkPNP I and AhPNP II on aldehyde-agarose (Serra et al. 2013b), were all loaded at mg g^{-1} quantities by almost three orders of magnitude less than that reported in this study. The main reason for this difference is likely related to the properties of the supports including the surface area, the functional group density, hydrophicity and so on. The particle size of MagEM is around 10 µm and the claimed functional group density is up to 3500 µmol g^{-1} (www.resynbio.com), while the common epoxide supports Sepabeads-EP are around 200 µm (www.resindion.com) and the density of epoxide groups on the surface of the beads is 600 µmol g^{-1} dry Sepabeads-EP (equivalent to that of Eupergit C) (Mateo et al. 2002; Katchalski-Katzir and Kraemer 2000). The claimed porous structure of the MagEM might provide suitable mass transfer characteristics as well as confer multipoint attachment for the stabilisation of multimeric proteins. The benefits of the high binding capacity supports the efficiency of biocatalytic reactions by reducing the cost (lower carrier cost) while further improving volumetric activity. The high functional group density, thus high binding capacities, reduces the risk of enzyme leaking which may lead to the product contamination, a critical problem in the product of pharmaceutical intermediates. With high binding capacity supports, we deduced that enzymes can be safely immobilized at lower enzyme loading and maintain a high activity. This was especially impressive for TtPyNP, which had a

residual activity of 150 % (Figure 4.1A). The reason might involve (i) a favourable partition effect (Illanes 2008) that the substrate (phosphate ion) was absorbed by the polymer support through ionic interaction; (ii) some beneficial structural change for the improvement of the enzyme activity. The specific cause warrants further investigation (e.g. determine for the stability, pH profile, and K_m & V_{max} of each preparation).

Due to the high volumetric activity, non-natural substrates with poor substrate properties were chosen for subsequent characterisation since the natural substrates were depleted too rapid under the optimal conditions. With this strategy, increased quantities of enzyme could be loaded according to the requirements of industrial application. Finally, the outstanding properties of the prepared biocatalysts were further verified in the synthetic application. Two 2,6-dihalogenated purine nucleosides, 26DCP-R and 6C2FP-R, as important precursors for the synthesis of many anti-leukaemia drugs (e.g. Cladribine, Clofarabine, Fludarabine) and other new nucleosides of potential pharmaceutical importance, were successfully synthesized and the product yields were optimized. The efficient recovery of the magnetic beads allowed for the step-by-step analysis study of the enzymatic cascade reactions. We found that PNP coupled with PyNP worked more efficiently than PNP alone with α-*D*-Rib-1P and purine base. In contrast to previous reports (Ubiali et al. 2004; Lewkowicz et al. 2000), very low concentration of phosphate (1.4 mM instead of 10-30 mM) was found as the optimal condition for the product formation in the one-pot reaction.

4.4. Conclusion

This work demonstrates that the proposed MagEM procedure is ideally suitable for the immobilization of multimeric thermozymes at the example of TtPyNP and GtPNP. Following a simple protocol, the resulting biocatalyst displays a combination of high enzyme activity maintenance with high enzyme loading, as well as improved enzyme stability. The magnetic property makes the recovery of the biocatalysts much easier at the scale employed in this study, while the separation at larger scale has yet to be determined.

The screening of the immobilization conditions suggests that binding buffer pH, enzyme loading, binding temperature and binding time are important factors to be considered. A non-linear model based on DoE was established to describe the relationship between an immobilization yield and binding time with binding temperature.

The application of the biocatalyst (4.6 mg mL^{-1}) in the synthesis of 26DCP-R and 6C2FP-R resulted product yields of 78.5 % and 85.5 % with productivities of 1.51 g L^{-1} h^{-1} and 1.97 g L^{-1} h^{-1}, respectively.

Enzyme immobilization with MagEM provided several advantages for this small scale study. However, we foresee several challenges in the separation of the magnetic microspheres at larger scale. Further magnetic supports potentially negate the use of a continuous process. Proposed methods of separation for large scale include an electromagnetic stick within the reactor, or an electromagnetic mesh, which might provide a solution for using magnetic beads at increased scale.

Chapter 5. Enzymatic synthesis (paper IV)

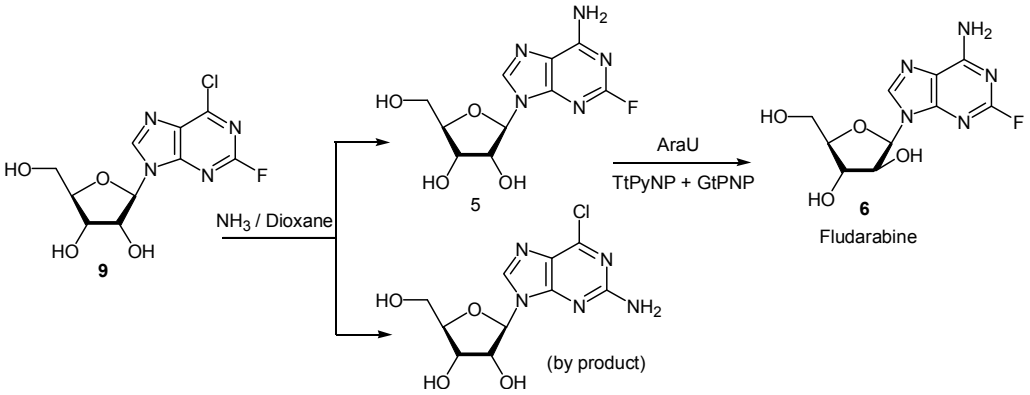
The foregoing investigations of 26DCP-(d)R and 6C2FP-(d)R synthesis (by the thermostable NPs) have been described in details in Chapter 3.1 & 3.2. However, all the experiments were performed in small scale (0.2 to 1 mL). Besides, low concentration of the substrate (1 mM purine base) was used.

In this work, the solubility of different purine bases was measured (paper IV, table 2). A clear picture of the temperature dependence of the base solubility can be drawn. That is, much higher solubility especial for 6C2FP was observed at the higher temperature, which implied that very high substrate concentration could be applied (up to 150.2 mM at 80 °C). On the other side, extremely low solubility was observed on 2-Cl-Ade (0.5 mM at 80 °C) as well as on 2-F-Ade (4.0 mM at 80 °C). The former is the purine base in Cladribine and Clofarabine; the latter in Fludarabine. These data suggest that although PNPs could catalyse the synthesis of Cladribine from 2-Cl-Ade efficiently (paper IV, table 4) the exceedingly low solubility of 2-Cl-Ade precludes its application. Therefore, the new idea is to use these highly soluble purine nucleosides to offer the halogenated purine base instead of using the insoluble base directly. To prove this idea, we need perform more experiments regards to the ammonolysis of dihalogenated purine nucleosides and transglycosylation of two nucleosides (Scheme 5.1 & 5.2).

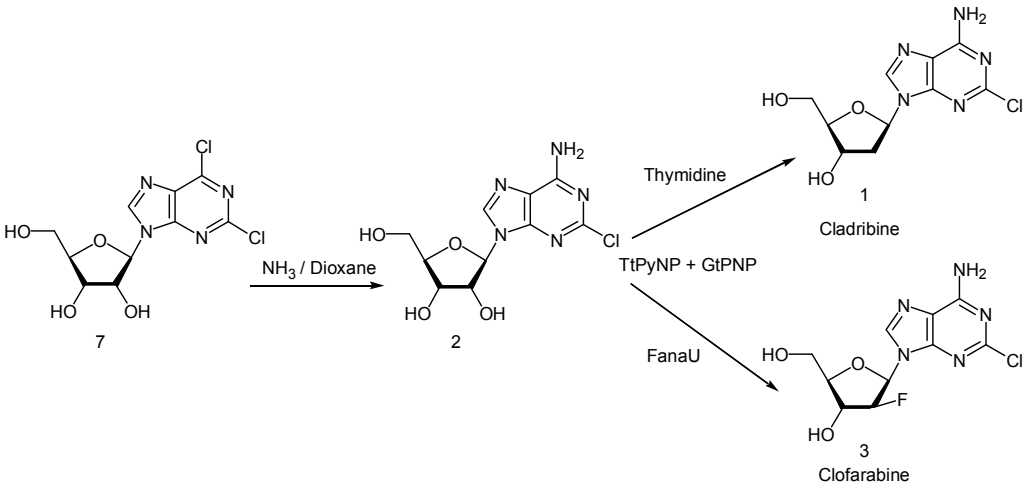
The used biocatalysts have intrinsic excellent thermostability and are suitable for industry application. The obtained compounds (26DCP-R and 6C2FP-R) implied versatile transformations to produce diverse nucleoside drugs, such as Cladribine **1**, Clofarabine **3**, and Fludarabine **6** (numbering according to paper IV, the same below). This approach is a breakthrough in the low solubility problem for the synthesis of 2Cl-Ade and 2F-Ade nucleosides.

The importance of 2,6-dihalogenated purine nucleosides (**7-10**) is not only confined to their potential bioactivities. Furthermore, they are precursors that can be converted to other well-known anti-viral or anti-tumour drugs by one-step ammonolysis. For example, 6C2FP-R (**9**) can be used to synthesize the cytotoxic compound 2-fluoroadenosine (**5**) by direct ammonolysis. With AraU as pentofuranosyl donor, **5** can be further enzymatically converted into another

antitumor drug Fludarabine **6** (Scheme 5.1). Here, the heterocyclic base is offered by 6C2FP-R not by 2F-Ade because the latter has extremely poor solubility while 6C2FP-R has a more than 150-time higher solubility than 2F-Ade at 40 °C in aqueous solution (212 mM *vs* 1.4 mM, paper IV, table 2). The ammonolysis product **5** is supposed to be water soluble and can provide its base by enzymatic transglycosylation.



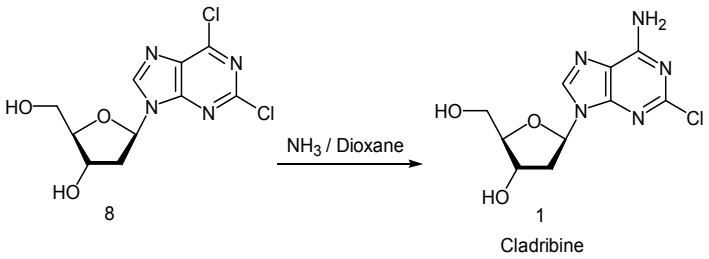
Scheme 5.1. 6C2FP-R (9) used for other nucleosides synthesis.



Scheme 5.2. 26DCP-R (7) used for other nucleosides synthesis.

In a similar way, 26DCP-R (**7**) can be converted to 2-chloroadenosine (**2**), and the latter can be used as 2Cl-Ade donor for the synthesis of at least two antitumor drugs (Scheme 5.2), Cladribine **1** and Clofarabine **3**. Both are used for the treatment of leukaemia.

26DCP-dR (**8**) can be converted to **1** by one-step ammonolysis (Scheme 5.3), which would be an alternative route for the synthesis of Cladribine. Komatsu and Araki reported that the ammonolysis of **8** performed in NH_4OH/CH_3CN resulted in **1** of 98 % HPLC yield, in which they used *E. coli* PNP to synthesize **8** (Komatsu and Araki 2005), however, **8** were converted from the highly unstable 2-deoxyribose-1-phosphate, which has to be prepared freshly.



Scheme 5.3. 26DCP-dR (8) converted to Cladribine (1).

Therefore, a great advantage of this method is that it enables the universal preparation of the modified sugar moiety and/or modified purine nucleosides.

Overall, the main results obtained from this work are:

- TtPyNP, GtPyNP and GtPNP were immobilized and used as biocatalysts in the synthesis of 26DCP-R and 6C2FP-R. After optimization, the yields (tested by HPLC) were increased from 54 % and 63 % (substrate: 1 mM base) to 78 % (16 mM base) and 85 % (50 mM base), respectively.
- 6C2FP-R was synthesized in 50 mL by the immobilized enzymes with high substrate concentration. Product was successfully isolated and purified. The structure was confirmed by ^{1}H-NMR, ^{13}C-NMR, and the high resolution MS.
- The apparent activity assay of PNPs towards purine bases was established, which is useful to determine the synthesis rate (from the reverse reaction of phosphorolysis) without using the expensive intermediate (d)R-1-P. Conversely, the standard assay can only measure the phosphorolysis rate (forward reaction).

Chapter 6. Conclusion and outlook

The aims of this work have been achieved, based on (i) the recombinant thermostable nucleoside phosphorylases expression in *E. coli*, (ii) enzyme characterisation and the related reaction processes analysing, (iii) enzyme immobilization, (iv) the use of the immobilized biocatalysts for the synthesis of modified purine nucleosides, and (v) product (6C2FP-R) purification and structure confirmation. Here, the main results, conclusions and outlook of this work will be summarized.

6.1. Enzyme expression and characterisation

- An analysis of the enzymatic synthesis of nucleosides allows us to conclude that the combined chemo-enzymatic method manifests an efficient, simple and "green" method; and the enzymes – nucleoside phosphorylases (NPs), play the crucial important role in both cascade transformation and transglycosylation, which prompts us to turn our attention to thermostable NPs.
- Five NPs from thermophilic microorganisms have been successfully expressed in *E. coli*, containing two pyrimidine nucleosides from *G. thermoglucosidasius* (GtPyNP) and *T. thermophilus* (TtPyNP); three PNPs from *D. geothermalis* (DgPNP), *G. thermoglucosidasius* (GtPNP) and *A. pernix* (ApMTAP).
- The purified enzyme productivity (volumetric yield) was calculated based on the OD_{600} measurement ($\propto$ cells number) and it is shown that the expression levels agree with the CAI prediction, which in the order of expression difficulty: ApMTAP > TtPyNP > DgPNP > GtPyNP > GtPNP.
- The factors which have big influence on the expression level are: his-tag position and length (matter of test), 5'mRNA stability (highly stable in the genes of thermozymes, which can be relieved by increasing cultivation temperature and silence mutation), disulfide bond (often appears in the hypothermophilies Archaea) and rare codon (difficult for *E. coli* to express). EnPresso medium (controlled substrate delivery technology) is effective for relieving the negative factors.

- The effect of temperature on the activity of each enzyme was tested from 30 °C to 99 °C. DgPNP, GtPNP, and GtPyNP showed an optimum temperature at 55 °C, 70 °C, and 60 °C, respectively. TtPyNP and ApMTAP showed a reaction rate increased with the temperature until the highest point tested (99 °C). However, a dramatic activity lost at low temperature was evident for TtPyNP and ApMTAP, which suggests that an activation temperature is needed for those thermozymes.
- All the NPs in this study exhibited expected thermostability according to their optimum temperatures. The effect of organic solvent was tested for PNPs and they showed resistance to organic solvent (MeOH and DMSO) in some degree. ApMTAP showed an unusual high stability which resists heat denaturation (90 °C) and even keeps active after running SDS-PAGE !
- TtPyNP and GtPNP displayed a particularly high catalytic activity (high k_{cat}) towards natural substrates comparing to the other reported NPs.
- TtPyNP showed a broad substrate specificity which can phosphorolyse 2'- or 3'-NH_2, 2'-F (*ribo*- or *arabino*-) and 2'-OH (*arabino*) substituted pyrimidine nucleosides, and also natural purine nucleosides as well.
- All the PNPs showed a rather broad substrate spectrum: modified on sugar moiety 2'- or 3'-NH_2, 2'-F (*ribo*- or *arabino*-) and 2'-OH (*arabino*); also on purine base 2,6-Cl or F substituted purine nucleosides can be accepted as substrate at different activity level.
- Amino acid sequence analysis by multiple sequence alignment provided the important information of the relationship between the enzyme substrate specificity and the key amino acid residues of the enzyme active sites. Through this, DgPyNP is suggested to be promising biocatalyst; "low-mm" PNPs from thermophiles are supposed to be tolerant to sugar moiety modified nucleosides.

6.2. Transglycosylation processes

Three types of translycosylation reactions by coupled PyNP and PNP were studied. With Urd as pentofuranosyl donor, an equilibrium product yield will be reached; with Thd (2'-deoxynucleoside) as donor, a drop of product yield due to the degradation of the intermediate compound dR-1-P was observed; with 2'-fluoro

substituted uridine as donor, reaction was kinetically controlled by the enzyme activity towards the substrates, for obtaining a high product yield, big enzyme amount and long reaction time are suggested.

6.3. Immobilization

- TtPyNP and GtPNP were successfully immobilized on the magnetic microsphere beads with high enzyme activity maintenance and high enzyme loading, as well as improved enzyme stability.
- The screening of the immobilization conditions suggests that binding buffer pH, enzyme loading, binding temperature and binding time are important factors to be considered.
- The application of the immobilized enzymes (4.6 mg mL^{-1}) in the synthesis of 26DCP-R and 6C2FP-R achieved product yields of 78.5 % and 85.5 % with respective productivity of 1.51 g L^{-1} h^{-1} and 1.97 g L^{-1} h^{-1}.

6.4. Enzymatic synthesis

- The solubility of different purine bases was measured. Increasing temperature can remarkably increase the 6C2FP solubility, which implied that very high substrate concentration could be applied (up to 150.2 mM at 80 °C). On the other side, extremely low solubility was observed on 2-Cl-Ade (0.5 mM at 80 °C) as well as on 2-F-Ade (4.0 mM at 80 °C).
- The immobilized enzymes were used in the synthesis of 26DCP-R and 6C2FP-R. After optimization, the product yields (HPLC) were increased from 54 % and 63 % (substrate: 1 mM base) to 78 % (16 mM base) and 85 % (50 mM base), respectively.
- 6C2FP-R was synthesized in 50 mL by the immobilized enzymes with high substrate concentration. Product was successfully isolated and purified by normal silica column. The structure was confirmed by 1H-NMR, 13C-NMR, and high resolution MS.

6.5. Outlook

The biocatalytic properties of the thermostable nucleoside phosphorylases in this study reveal their great practical potential. To attain the goal for the application, future work could include:

- Comparing immobilization on magnetic microspheres and other possibilities for the large scale biotransformation.
- Establishing a continuous process in lab scale.
- Investigation of enzyme cascade process (including ribokinase and phosphopentomutase or the enzymes which could shift transglycosylation equilibrium).
- Development of kinetic model for cascade enzymatic transformation.
- Optimization of enzyme process by combination of design of experiments (DoE) and modelling.
- Investigation of novel method for product isolation and purification (ion-exchange resins coupled with reaction process engineering).
- Searching for new biocatalysts from diverse sources through rational bioinformatics analysing and data mining (e.g. DgPyNP, thermostable NDTs).

Chapter 7. Appendix

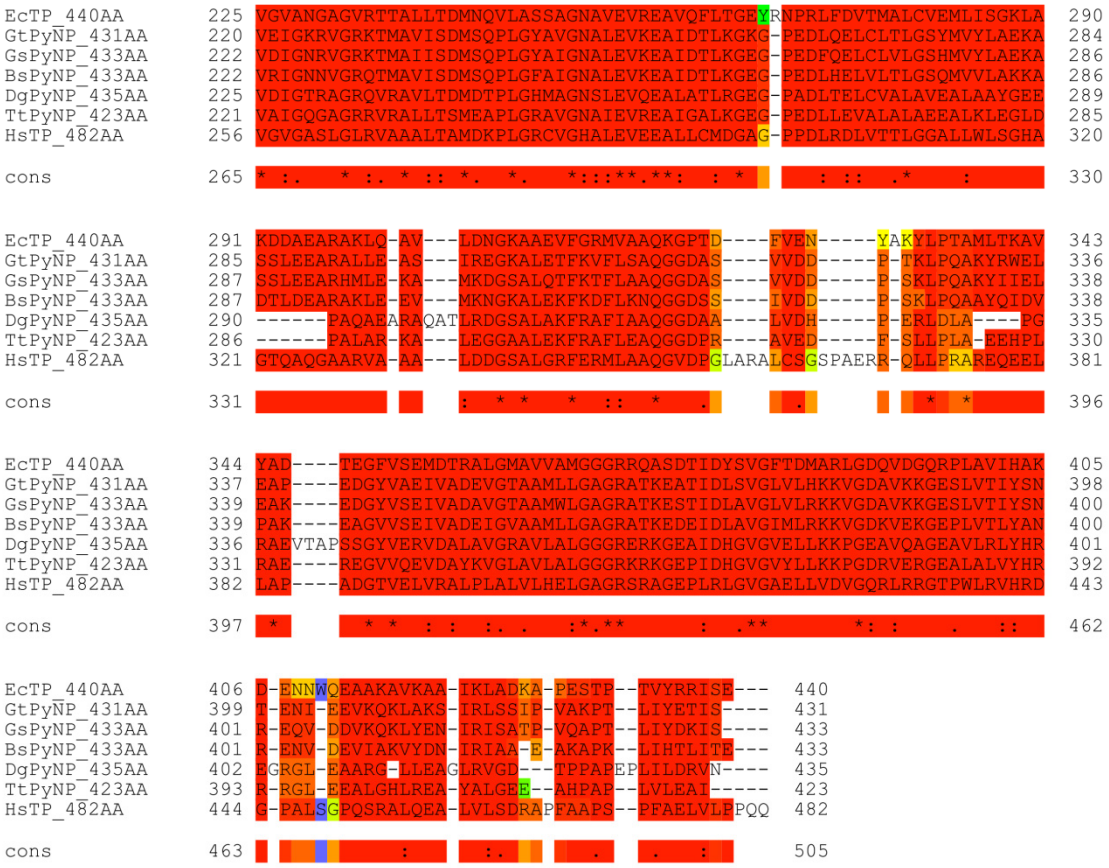
7.1. Multiple sequence alignment

7.1.1. Original alignment result of PyNP & TP from 3D-Coffee

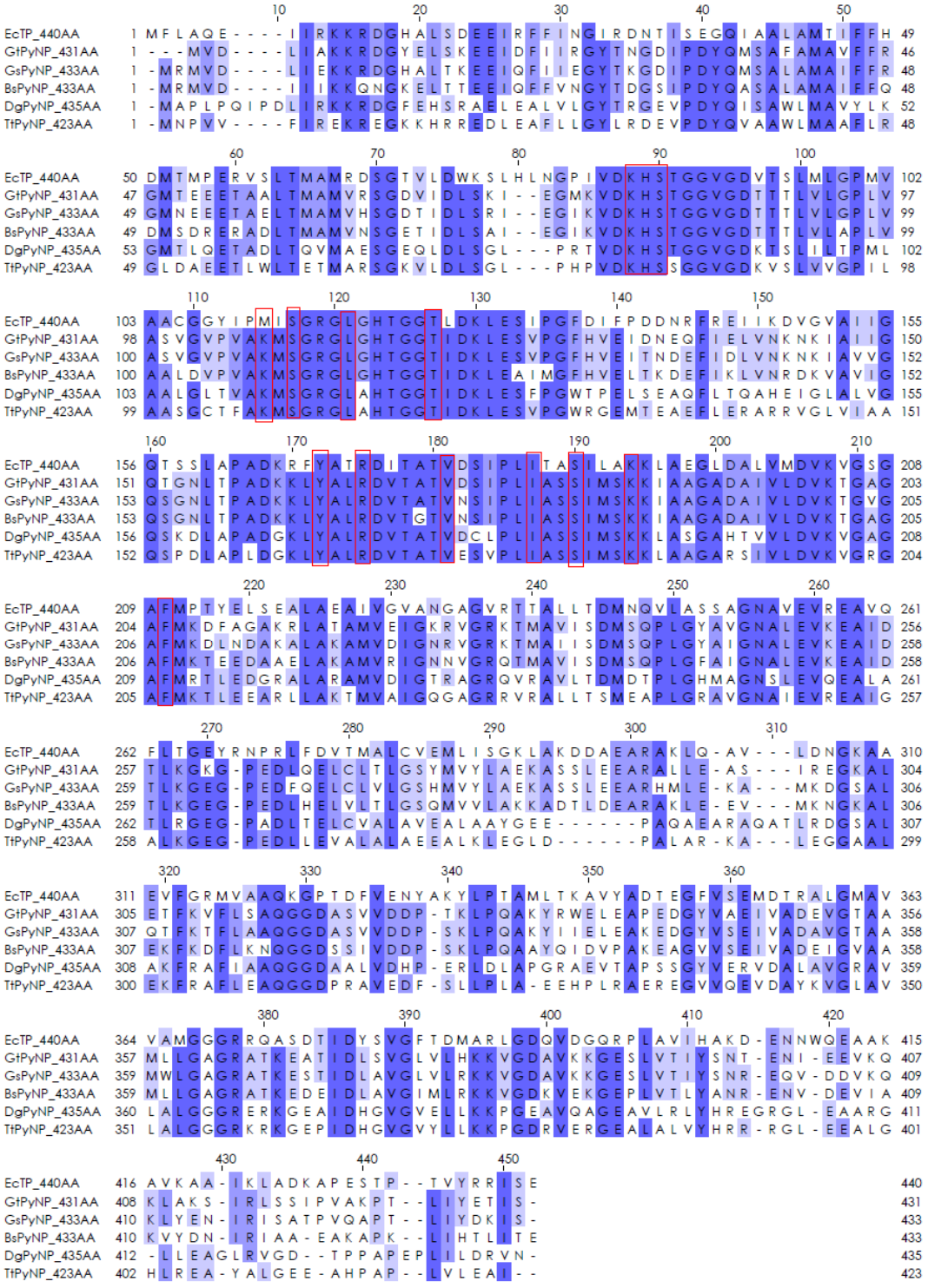
Expresso alignment result

MSA The multiple sequence alignment result as produced by T-coffee.

```
T-COFFEE, Version_10.00.r1613 (2013-10-22 15:49:09 - Revision 1613 - Build 432)
Cedric Notredame
SCORE=98
*
 BAD  AVG  GOOD
*
EcTP_440AA      :  98
GtPyNP_431AA    :  99
GsPyNP_433AA    :  98
BsPyNP_433AA    :  98
DgPyNP_435AA    :  98
TtPyNP_423AA    :  98
HsTP_482AA      :  98
cons            :  98

EcTP_440AA        1 --------------------------------MFLAQE----IIRKKRDGHALSDEEIRFFINGIR   30
GtPyNP_431AA      1 -----------------------------------MVD----LIAKKRDGYELSKEEIDFIIRGYT   27
GsPyNP_433AA      1 --------------------------------MRMVD----LIEKKRDGHALTKEEIQFIIEGYT   29
BsPyNP_433AA      1 --------------------------------MRMVD----IIIKKQNGKELTTEEIQFFVNGYT   29
DgPyNP_435AA      1 --------------------------------MAPLPQIPDLIRKKRDGFEHSRAELEALVLGYT   33
TtPyNP_423AA      1 --------------------------------MNPVV----FIREKREGKKHRREDLEAFLLGYL   29
HsTP_482AA        1 MAALMTPGTGAPPAPGDFSGEGSQGLPDPSPEPKQLPE----LIRMKRDGGRLSEADIRGFVAAVV   62

cons              1                                    :*  *::*      ::  :: .        66


EcTP_440AA       31 DNTISEGQIAALAMTIFFHDMTMPERVSLTMAMRDSGTVLDWKS--LHLNGPIVDKHSTGGVGDVT   94
GtPyNP_431AA     28 NGDIPDYQMSAFAMAVFFRGMTEEETAALTMAMVRSGDVIDLSK--I--EGMKVDKHSTGGVGDTT   89
GsPyNP_433AA     30 KGDIPDYQMSALAMAIFFRGMNEEETAELTMAMVHSGDTIDLSR--I--EGIKVDKHSTGGVGDTT   91
BsPyNP_433AA     30 DGSIPDYQASALAMAIFFQDMSDRERADLTMAMVNSGETIDLSA--I--EGIKVDKHSTGGVGDTT   91
DgPyNP_435AA     34 RGEVPDYQISAWLMAVYLKGMTLQETADLTQVMAESGEQLDLSG--L---PRTVDKHSTGGVGDKT   94
TtPyNP_423AA     30 RDEVPDYQVAAWLMAAFLRGLDAEETLWLTETMARSGKVLDLSG--L---PHPVDKHSSGGVGDKV   90
HsTP_482AA       63 NGSAQGAQIGAMLMAIRLRGMDLEETSVLTQALAQSGQQLEW-PEAW--RQQLVDKHSTGGVGDKV  125

cons             67 .     * .*  *:  ::.:   *   ** .:  **  ::     *****:***** .        132


EcTP_440AA       95 SLMLGPMVAACGGYIPMISGRGLGHTGGTLDKLESIPGFDIFP--DDNRFREIIKDVGVAIIGQTS  158
GtPyNP_431AA     90 TLVLGPLVASVGVPVAKMSGRGLGHTGGTIDKLESVPGFHVEI--DNEQFIELVNKNKIAIIGQTG  153
GsPyNP_433AA     92 TLVLGPLVASVGVPVAKMSGRGLGHTGGTIDKLESVPGFHVEI--TNDEFIDLVNKNKIAVVGQSG  155
BsPyNP_433AA     92 TLVLAPLVAALDVPVAKMSGRGLGHTGGTIDKLEAIMGFHVELTKDE--FIKLVNRDKVAVIGQSG  155
DgPyNP_435AA     95 SLILTPMLAALGLTVAKMSGRGLAHTGGTIDKLESFPGWTPEL--SEAQFLTQAHEIGLALVGQSK  158
TtPyNP_423AA     91 SLVVGPILAASGCTFAKMSGRGLAHTGGTIDKLESVPGWRGEM--TEAEFLERARRVGLVIAAQSP  154
HsTP_482AA      126 SLVLAPALAACGCKVPMISGRGLGHTGGTLDKLESIPGFNVIQ--SPEQMQVLLDQAGCCIVGQSE  189

cons            133 :*:: * :*: .  .. :*****.*****:****:. *:      :          : .*:     198


EcTP_440AA      159 SLAPADKRFYATRDITATVDSIPLITASILAKKLAEGLDALVMDVKVGSGAFMPTYELSEALAEAI  224
GtPyNP_431AA    154 NLTPADKKLYALRDVTATVDSIPLIASSIMSKKIAAGADAIVLDVKTGAGAFMKDFAGAKRLATAM  219
GsPyNP_433AA    156 NLTPADKKLYALRDVTATVNSIPLIASSIMSKKIAAGADAIVLDVKTGVGAFMKDLNDAKALAKAM  221
BsPyNP_433AA    156 NLTPADKKLYALRDVTGTVNSIPLIASSIMSKKIAAGADAIVLDVKTGAGAFMKTEEDAAELAKAM  221
DgPyNP_435AA    159 DLAPADGKLYALRDVTATVDCLPLIASSIMSKKLASGAHTVVLDVKVGAGAFMRTLEDGRALARAM  224
TtPyNP_423AA    155 DLAPLDGKLYALRDVTATVESVPLIASSIMSKKLAAGARSIVLDVKVGRGAFMKTLEEARLLAKTM  220
HsTP_482AA      190 QLVPADGILYAARDVTATVDSLPLITASILSKKLVEGLSALVVDVKFGGAAVFPNQEQARELAKTL  255

cons            199 .*.* *  :** **:*.**:.:***::**::**:. *  ::*:*** * .*.:     .  ** ::  264
```

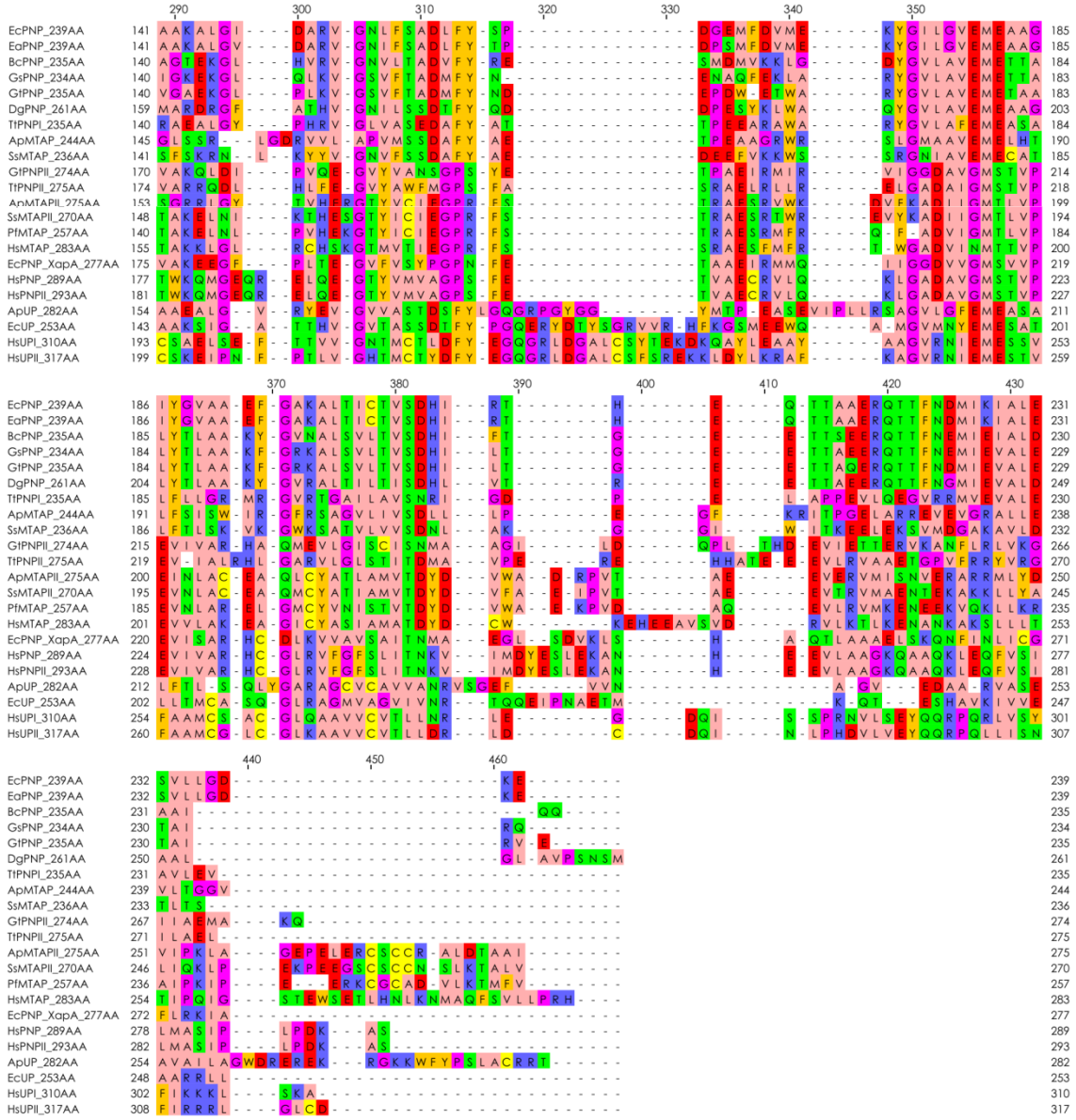

EcTP_440AA 225 VGVANGAGVRTTALLTDMNQVLASSAGNAVEVREAVQFLTGEYRNPRLFDVTMALCVEMLISGKLA 290
GtPyNP_431AA 220 VEIGKRVGRKTMAVISDMSQPLGYAVGNALEVKEAIDTLKGKG-PEDLQELCLTLGSYMVYLAEKA 284
GsPyNP_433AA 222 VDIGNRVGRKTMAIISDMSQPLGYAIGNALEVKEAIDTLKGEG-PEDFQELCLVLGSHMVYLAEKA 286
BsPyNP_433AA 222 VRIGNNVGRQTMAVISDMSQPLGFAIGNALEVKEAIDTLKGEG-PEDLHELVLTLGSQMVVLAKKA 286
DgPyNP_435AA 225 VDIGTRAGRQVRAVLTDMDTPLGHMAGNSLEVQEALATLRGEG-PADLTELCVALAVEALAAYGEE 289
TtPyNP_423AA 221 VAIGQGAGRRVRALLTSMEAPLGRAVGNAIEVREAIGALKGEG-PEDLLEVALALAEEALKLEGLD 285
HsTP_482AA 256 VGVGASLGLRVAAALTAMDKPLGRCVGHALEVEEALLCMDGAG-PPDLRDLVTTLGGALLWLSGHA 320
cons 265 330
EcTP_440AA 291 KDDAEARAKLQ-AV---LDNGKAAEVFGRMVAAQKGPTD----FVEN-----YAKYLPTAMLTKAV 343
GtPyNP_431AA 285 SSLEEARALLE-AS---IREGKALETFKVFLSAQGGDAS----VVDD-----P-TKLPQAKYRWEL 336
GsPyNP_433AA 287 SSLEEARHMLE-KA---MKDGSALQTFKTFLAAQGGDAS----VVDD-----P-SKLPQAKYIIEL 338
BsPyNP_433AA 287 DTLDEARAKLE-EV---MKNGKALEKFKDFLKNQGGDSS----IVDD-----P-SKLPQAAYQIDV 338
DgPyNP_435AA 290 ------PAQAEARAQATLRDGSALAKFRAFIAAQGGDAA----LVDH-----P-ERLDLA----PG 335
TtPyNP_423AA 286 ------PALAR-KA---LEGGAALEKFRAFLEAQGGDPR----AVED-----F-SLLPLA-EEHPL 330
HsTP_482AA 321 GTQAQGAARVA-AA---LDDGSALGRFERMLAAQGVDPGLARALCSGSPAERR-QLLPRAREQEEL 381
cons 331 396
EcTP_440AA 344 YAD----TEGFVSEMDTRALGMAVVAMGGGRRQASDTIDYSVGFTDMARLGDQVDGQRPLAVIHAK 405
GtPyNP_431AA 337 EAP----EDGYVAEIVADEVGTAAMLLGAGRATKEATIDLSVGLVLHKKVGDAVKKGESLVTIYSN 398
GsPyNP_433AA 339 EAK----EDGYVSEIVADAVGTAAMWLGAGRATKESTIDLAVGLVLRKKVGDAVKKGESLVTIYSN 400
BsPyNP_433AA 339 PAK----EAGVVSEIVADEIGVAAMLLGAGRATKEDEIDLAVGIMLRKKVGDKVEKGEPLVTLYAN 400
DgPyNP_435AA 336 RAEVTAPSSGYVERVDALAVGRAVLALGGGRERKGEAIDHGVGVELLKKPGEAVQAGEAVLRLYHR 401
TtPyNP_423AA 331 RAE----REGVVQEVDAYKVGLAVLALGGGRKRKGEPIDHGVGVYLLKKPGDRVERGEALALVYHR 392
HsTP_482AA 382 LAP----ADGTVELVRALPLALVLHELGAGRSRAGEPLRLGVGAELLVDVGQRLRRGTPWLRVHRD 443
cons 397 462
EcTP_440AA 406 D-ENNWQEAAKAVKAA-IKLADKA-PESTP--TVYRRISE--- 440
GtPyNP_431AA 399 T-ENI-EEVKQKLAKS-IRLSSIP-VAKPT--LIYETIS---- 431
GsPyNP_433AA 401 R-EQV-DDVKQKLYEN-IRISATP-VQAPT--LIYDKIS---- 433
BsPyNP_433AA 401 R-ENV-DEVIAKVYDN-IRIAA-E-AKAPK--LIHTLITE--- 433
DgPyNP_435AA 402 EGRGL-EAARG-LLEAGLRVGD---TPPAPEPLILDRVN---- 435
TtPyNP_423AA 393 R-RGL-EEALGHLREA-YALGEE--AHPAP--LVLEAI----- 423
HsTP_482AA 444 G-PALSGPQSRALQEA-LVLSDRAPFAAPS--PFAELVLPPQQ 482
cons 463 505

7.1.2. Recoloured alignment of PyNP and TP



The shading represents the degree of sequence identity. The active sites are indicated in red box (Pugmire and Ealick 1998; Mendieta et al. 2004).

7.1.3. Recoloured alignment of PNP (I/II) and UP

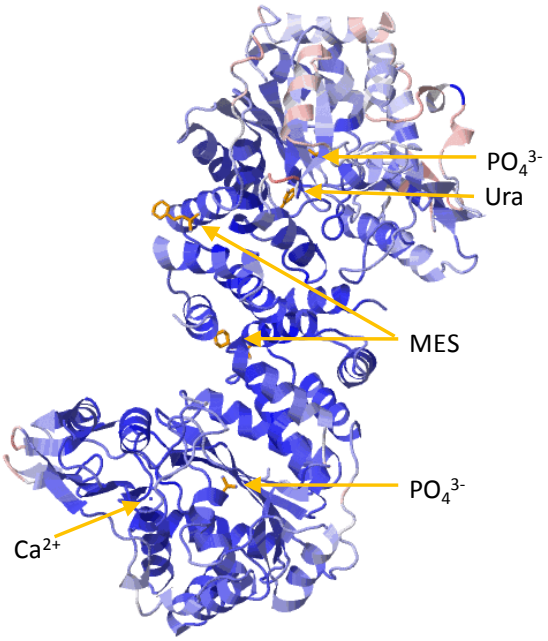


The amino acid residues are coloured according to their physicochemical properties as follows:

Aliphatic/hydrophobic	ILVAM
Aromatic	FWY
Positive	KRH
Negative	DE
Hydrophilic	STNQ
Conformationally special	PG
Cysteine	C

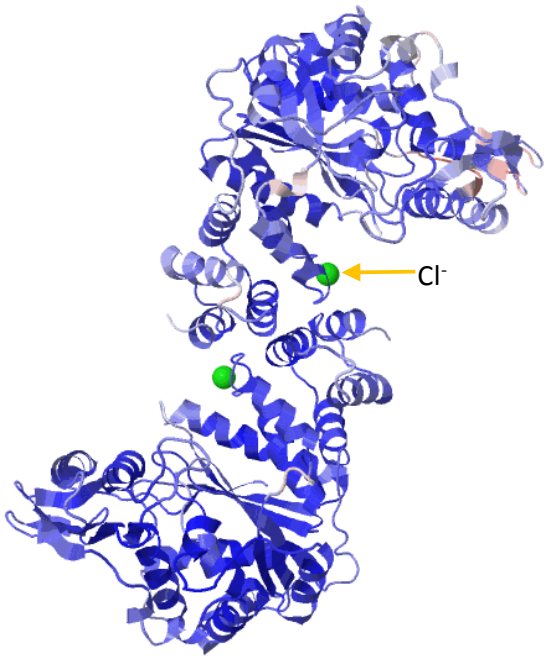
7.2. 3D-structure of NPs by homology modelling

7.2.1. Structures of PyNPs

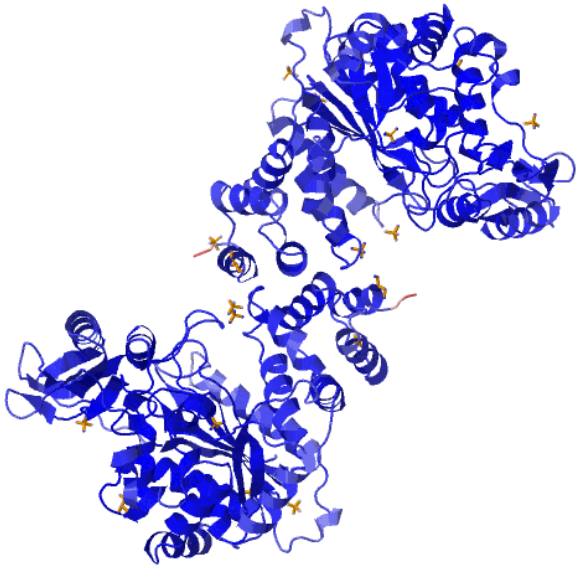


GtPyNP (PDB:1BRW; GsPyNP)

MES: 2-(n-morpholino)-ethanesulfonic acid

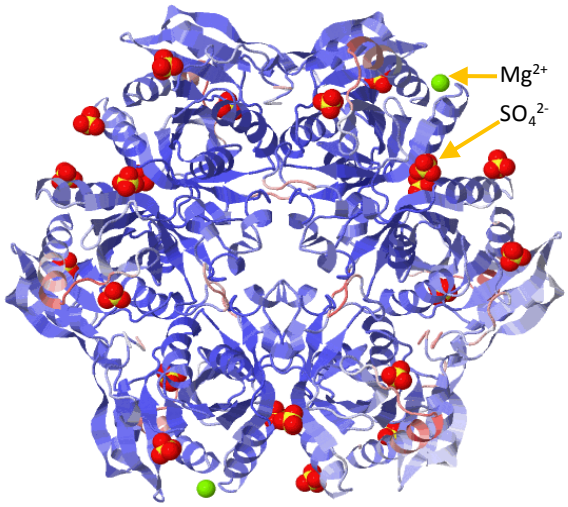


TtPyNP (PDB:2DSJ)



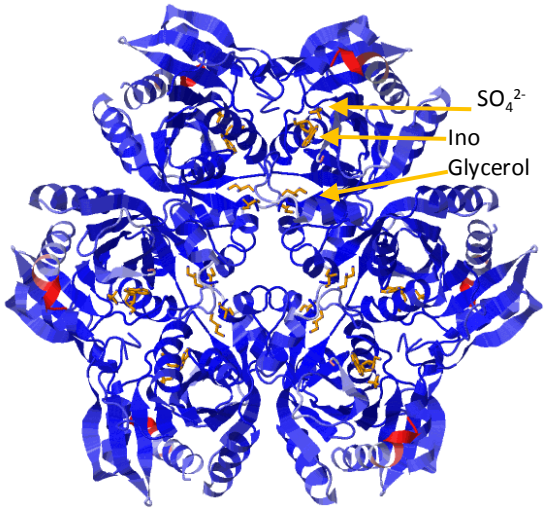
EcTP (PDB:4eaf)

7.2.2. Structures of PNPs



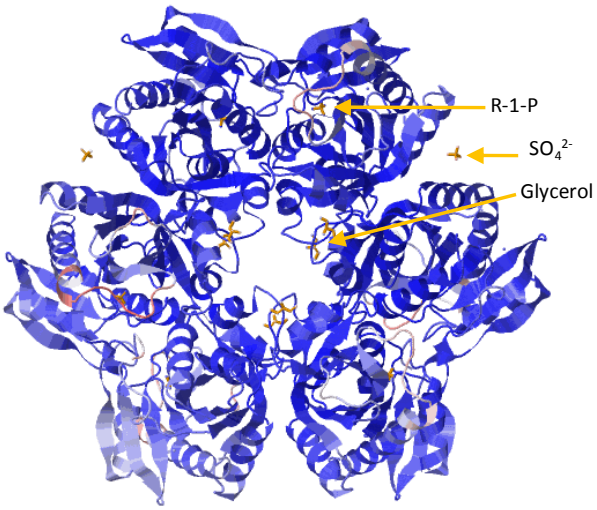
DgPNP

(PDB: 4MAR; MrPNP; contains 15 selenomethionines)

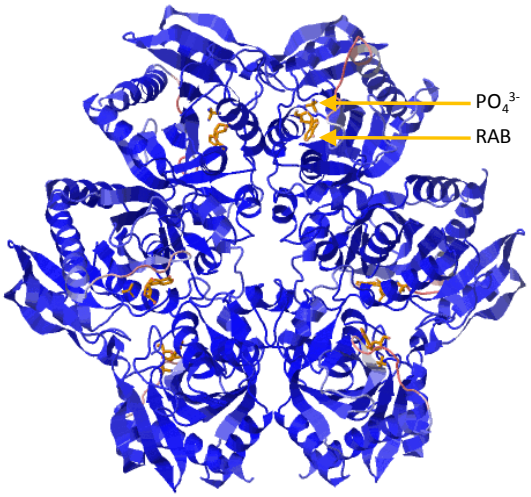


GtPNP

(PDB: 3UAX; BcPNP)



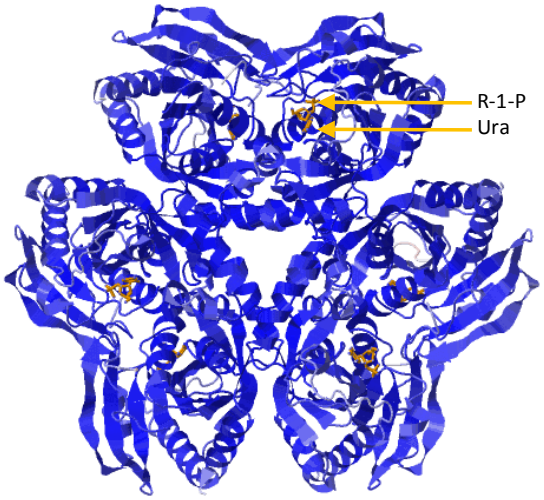
TtPNPI (PDB: 1ODL)



EcPNP (PDB:1PW7)

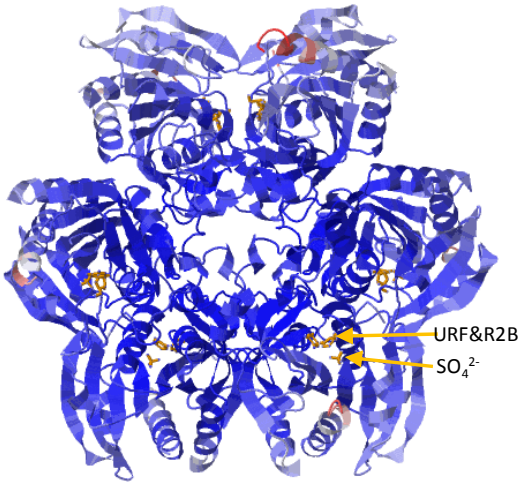
RAB: 2-(6-amino-purin-9-yl)-5-hydroxymethyl-tetrahydro-furan-3,4-diol

7.2.3. Structures of UPs



ApUP

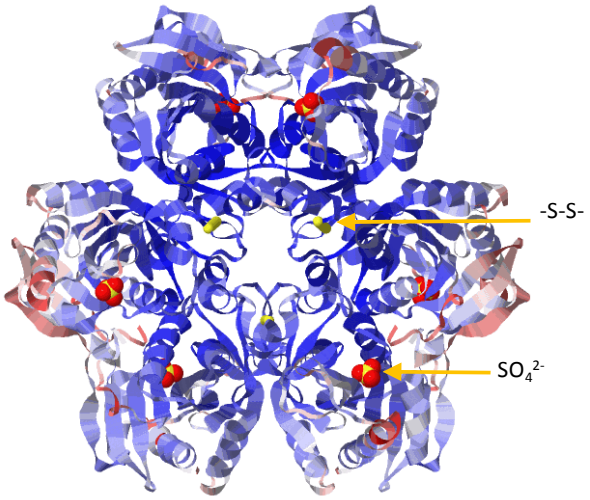
(PDB:3QPB; SpUP)



EcUP (PDB: 3KVV)

URF: 5-fluorouracil; R2B: 1,4-anhydro-D-erythro-pent-1-enitol

7.2.4. Structure of MTAP



ApMTAP

(PDB: 1JPV; SsMTAP)

7.3. Supporting information (paper III)

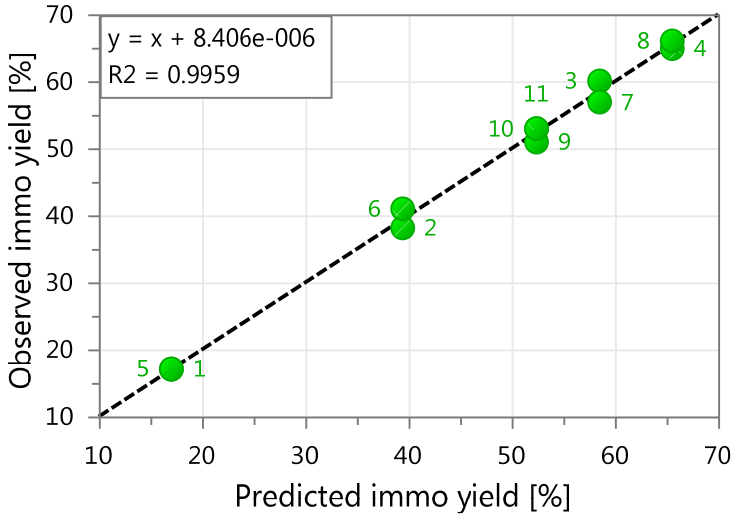


Fig. S1. Observed vs predicted plot for immobilization yield.

Table S1. Analysis of variance (ANOVA) for the Immo_Yield model.

Immo_Yield	DF	SS	MS (variance)	F	p	SD
Total	11	27352	2486.55			
Constant	1	24393.1	24393.1			
Total corrected	10	2958.91	295.891			17.2015
Regression	4	2946.74	736.686	363.296	0.000	27.142
Residual	6	12.1667	2.02778			1.424
Lack of Fit (Model error)	4	9.50003	2.37501	1.78125	0.390	1.54111
Pure error (Replicate error)	2	2.66667	1.33334			1.1547
	N = 11	Q2 =	0.985	Cond. no. =	3.601	
	DF = 6	R2 =	0.996	RSD =	1.424	
		R2 adj. =	0.993			

ANOVA, Statistical appendix (MODDE 10)

Analysis of variance (ANOVA) partitions the total variation of a selected response SS (Sum of Squares corrected for the mean) into a part due to the regression model and a part due to the residuals.

$SS_{TotC} = SS_{regr} + SS_{resid}$

If there are replicated observations (experiments), the residual sum of squares is further partitioned into pure error (SS_{pe}) and Lack of fit (SS_{lof}).

$SS_{resid} = SS_{pe} + SS_{lof}$

$DF_{resid} = (n - p)$

$SS_{pe} = \sum (e_{ki} - e_k)^2$

$DF_{pe} = \sum (n_k - 1)$

$DF_{lof} = n - p - \sum (n_k - 1)$

where

$\sum$ loops over ki resp k.

n = number of experimental runs (excluding missing values).

$\mathbf{n_k}$ = number of replicates in the k^{th} set.

p = number of terms in the model, including the constant.

$\mathbf{e_k}$ = average of the n_k residuals in the k^{th} set of replicates.

References

Almendros, M., Berenguer, J., Sinisterra, J.-V., **2012**. Thermus thermophilus Nucleoside Phosphorylases Active in the Synthesis of Nucleoside Analogues. Applied and Environmental Microbiology. 78, 3128-3135.

Almendros, M., Gago, J.-V. S., Carlos, J. B., **2009**. Thermus thermophilus Strains Active in Purine Nucleoside Synthesis. Molecules. 14, 1279-1287.

Anderson, B. G., Bauta, W. E., Cantrell, J. W. R., Engles, T., Lovett, D. P., **2008**. Isolation, Synthesis, and Characterization of Impurities and Degradants from the Clofarabine Process. Organic Process Research & Development. 12, 1229-1237.

Appleby, T. C., Mathews, I. I., Porcelli, M., Cacciapuoti, G., Ealick, S. E., **2001**. Three-dimensional Structure of a Hyperthermophilic 5′-Deoxy-5′-methylthioadenosine Phosphorylase from Sulfolobus solfataricus. Journal of Biological Chemistry. 276, 39232-39242.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., Notredame, C., **2006**. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Research. 34, W604-W608.

Averhoff, B., Müller, V., **2010**. Exploring research frontiers in microbiology: recent advances in halophilic and thermophilic extremophiles. Research in Microbiology. 161, 506-514.

Böttcher, T., Sieber, S. A., **2010**. Showdomycin as a Versatile Chemical Tool for the Detection of Pathogenesis-Associated Enzymes in Bacteria. Journal of the American Chemical Society. 132, 6964-6972.

Barai, V. N., Zinchenko, A. I., Eroshevskaya, L. A., Kalinichenko, E. N., Kulak, T. I., Mikhailopulo, I. A., **2002**. A universal biocatalyst for the preparation of base- and sugar-modified nucleosides via an enzymatic transglycosylation. Helvetica Chimica Acta. 85, 1901-1908.

Betbeder, D., Hutchinson, D. W., Richards, A. O. L., **1989**. The stereoselective enzymatic synthesis of 9-β-d-2′-deoxyribofuranosyl 1-deazapurine. Nucleic Acids Research. 17, 4217-4222.

Blidner, R. A., Hammer, R. P., Lopez, M. J., Robinson, S. O., Monroe, W. T., **2007**. Fully 2′-Deoxy-2′-Fluoro Substituted Nucleic Acids Induce RNA Interference in Mammalian Cell Culture. Chemical Biology & Drug Design. 70, 113-122.

Brady, D., Jordaan, J., **2009**. Advances in enzyme immobilisation. Biotechnol Lett. 31, 1639-50.

Bruins, M., Janssen, A., Boom, R., **2001**. Thermozymes and their applications. Applied Biochemistry and Biotechnology. 90, 155-186.

Bzowska, A., Kulikowska, E., Shugar, D., **2000**. Purine nucleoside phosphorylases: properties, functions, and clinical aspects. Pharmacol Ther. 88, 349-425.

Cacciapuoti, G., Bertoldo, C., Brio, A., Zappia, V., Porcelli, M., **2003**. Purification and characterization of 5'-methylthioadenosine phosphorylase from the hyperthermophilic archaeon Pyrococcus furiosus-substrate specificity and primary structure analysis. Extremophiles. 7, 159-168.

Cacciapuoti, G., Forte, S., Moretti, M. A., Brio, A., Zappia, V., Porcelli, M., **2005**. A novel hyperthermostable 5'-deoxy-5'-methylthioadenosine phosphorylase from the archaeon Sulfolobus solfataricus. FEBS J. 272, 1886-99.

Cacciapuoti, G., Gorassini, S., Mazzeo, M. F., Siciliano, R. A., Carbone, V., Zappia, V., Porcelli, M., **2007**. Biochemical and structural characterization of mammalian-like purine nucleoside phosphorylase from the Archaeon Pyrococcus furiosus. FEBS J. 274, 2482-95.

Cacciapuoti, G., Marabotti, A., Fuccio, F., Porcelli, M., **2011**. Unraveling the structural and functional differences between purine nucleoside phosphorylase and 5′-deoxy-5′-methylthioadenosine phosphorylase from the archaeon Pyrococcus furiosus. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics. 1814, 1358-1366.

Cacciapuoti, G., Moretti, M. A., Forte, S., Brio, A., Camardella, L., Zappia, V., Porcelli, M., **2004**. Methylthioadenosine phosphorylase from the archaeon Pyrococcus furiosus. European Journal of Biochemistry. 271, 4834-4844.

Cacciapuoti, G., Porcelli, M., Bertoldo, C., De Rosa, M., Zappia, V., **1994**. Purification and characterization of extremely thermophilic and thermostable 5'-methylthioadenosine phosphorylase from the archaeon Sulfolobus solfataricus. Purine nucleoside phosphorylase activity and evidence for intersubunit disulfide bonds. Journal of Biological Chemistry. 269, 24762-9.

Cacciapuoti, G., Servillo, L., Moretti, M., Porcelli, M., **2001**. Conformational changes and stabilization induced by phosphate binding to 5'-methylthioadenosine phosphorylase from the thermophilic archaeon Sulfolobus solfataricus. Extremophiles. 5, 295-302.

Caradoc-Davies, T. T., Cutfield, S. M., Lamont, I. L., Cutfield, J. F., **2004**. Crystal Structures of Escherichia coli Uridine Phosphorylase in Two Native and Three Complexed Forms Reveal Basis of Substrate Specificity, Induced Conformational Changes and Influence of Potassium. Journal of Molecular Biology. 337, 337-354.

Carson, D. A., Wasson, D. B., **1988**. Synthesis of 2',3'-Dideoxynucleosides by Enzymatic Trans-Glycosylation. Biochemical and Biophysical Research Communications. 155, 829-834.

Cen, Y., Sauve, A. A., **2010**. Efficient Syntheses of Clofarabine and Gemcitabine From 2-Deoxyribonolactone. Nucleosides, Nucleotides and Nucleic Acids. 29, 113-122.

Chuvikovsky, D. V., Esipov, R. S., Skoblov, Y. S., Chupova, L. A., Muravyova, T. I., Miroshnikov, A. I., Lapinjoki, S., Mikhailopulo, I. A., **2006**. Ribokinase from E. coli: Expression, purification, and substrate specificity. Bioorganic & Medicinal Chemistry. 14, 6327-6332.

Cleland, W. W., Steady State Kinetics. In: P. D. Boyer, (Ed.), The Enzymes: Kinetics and Mechanism, Vol. Volume II. Academic Press, California, **1970**, pp. 52-59.

De Clercq, E., **2009**. The history of antiretrovirals: key discoveries over the past 25 years. Reviews in Medical Virology. 19, 287-299.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., Gascuel, O., **2008**. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research. 36, W465-W469.

Desjardins, P., Hansen, J. B., Allen, M., Microvolume Spectrophotometric and Fluorometric Determination of Protein Concentration. Current Protocols in Protein Science. John Wiley & Sons, Inc., **2001**.

Dessanti, P., Zhang, Y., Allegrini, S., Tozzi, M. G., Sgarrella, F., Ealick, S. E., **2012**. Structural basis of the substrate specificity of Bacillus cereus adenosine phosphorylase. Acta Crystallogr D Biol Crystallogr. 68, 239-48.

Eisenmesser, E. Z., Bosco, D. A., Akke, M., Kern, D., **2002**. Enzyme Dynamics During Catalysis. Science. 295, 1520-1523.

Fernandez-Lafuente, R., **2009**. Stabilization of multimeric enzymes: Strategies to prevent subunit dissociation. Enzyme and Microbial Technology. 45, 405-418.

Fernandez-Lucas, J., Acebal, C., Sinisterra, J. V., Arroyo, M., de la Mata, I., **2010**. Lactobacillus reuteri 2 '-Deoxyribosyltransferase, a Novel Biocatalyst for Tailoring of Nucleosides. Applied and Environmental Microbiology. 76, 1462-1470.

Ferreira, A. C., Nobre, M. F., Rainey, F. A., Silva, M. T., Wait, R., Burghardt, J., Chung, A. P., da Costa, M. S., **1997**. Deinococcus geothermalis sp. nov. and Deinococcus murrayi sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. International Journal of Systematic Bacteriology. 47, 939-47.

Field, H. J., De Clercq, E., **2004**. Antiviral Drugs - a short history of their discovery and development. Microbiology Today. 31, 58-61.

Fresco-Taboada, A., de la Mata, I., Arroyo, M., Fernandez-Lucas, J., **2013**. New insights on nucleoside 2 '-deoxyribosyltransferases: a versatile Biocatalyst for one-pot one-step synthesis of nucleoside analogs. Applied Microbiology and Biotechnology. 97, 3773-3785.

Gerdes, K., **1988**. The Parb (Hok Sok) Locus of Plasmid-R1 - a General-Purpose Plasmid Stabilization System. Bio-Technology. 6, 1402-1405.

Hamamoto, T., Noguchi, T., Midorikawa, Y., **1996**. Purification and characterization of purine nucleoside phosphorylase and pyrimidine nucleoside phosphorylase from Bacillus stearothermophilus TH 6-2. Bioscience Biotechnology and Biochemistry. 60, 1179-1180.

Hatahet, F., Nguyen, V. D., Salo, K. E., Ruddock, L. W., **2010**. Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of E. coli. Microbial Cell Factories. 9, 67.

Hei, D. J., Clark, D. S., **1994**. Pressure Stabilization of Proteins from Extreme Thermophiles. Applied and Environmental Microbiology. 60, 932-939.

Henne, A., Bruggemann, H., Raasch, C., Wiezer, A., Hartsch, T., Liesegang, H., Johann, A., Lienard, T., Gohl, O., Martinez-Arias, R., Jacobi, C., Starkuviene, V., Schlenczeck, S., Dencker, S., Huber, R., Klenk, H. P., Kramer, W., Merkl, R., Gottschalk, G., Fritz, H. J., **2004**. The genome sequence of the extreme thermophile Thermus thermophilus. Nature Biotechnology. 22, 547-553.

Holguin, J., Cardinaud, R., **1975**. Trans-N-Deoxyribosylase - Purification by Affinity Chromatography and Characterization. European Journal of Biochemistry. 54, 505-514.

Holm, L., Sander, C., **1996**. Mapping the Protein Universe. Science. 273, 595-602.

Hori, N, Watanabe, M, Yamazaki, Y, Mikami, **1989**. Purification and characterization of second thermostable purine nucleoside phosphorylase in Bacillus stearothermophilus JTS 859. Agricultural and biological chemistry 53, 3219-3224.

Hori, N., Watanabe, M., Mikami, Y., **1991a**. The Effects of Organic Solvent on the Ribosyl Transfer Reaction by Thermostable Purine Nucleoside Phosphorylase and Pyrimidine Nucleoside Phosphorylase from Bacillus stearothermophilus JTS 859. Biocatalysis and Biotransformation. 4, 297-304.

Hori, N., Watanabe, M., Sunagawa, K., Uehara, K., Mikami, Y., **1991b**. Production of 5-methyluridine by immobilized thermostable purine nucleoside phosphorylase and

pyrimidine nucleoside phosphorylase from Bacillus stearothermophilus JTS 859. J Biotechnol. 17, 121-31.

Hori, N., Watanabe, M., Yamazaki, Y., Mikami, Y., **1990**. Purification and Characterization of Thermostable Pyrimidine Nucleoside Phosphorylase from *Bacillus stearothermophilus* JTS 859. Agricultural and Biological Chemistry. 54, 763-768.

Howell, H. G., Brodfuehrer, P. R., Brundidge, S. P., Benigni, D. A., Sapino, C., **1988**. Antiviral Nucleosides - a Stereospecific, Total Synthesis of 2'-Fluoro-2'-Deoxy-Beta-D-Arabinofuranosyl Nucleosides. Journal of Organic Chemistry. 53, 85-88.

Hutchinson, D. W., **1990**. New approaches to the synthesis of antiviral nucleosides. Trends in biotechnology. 8, 348-353.

Illanes, A., **2008**. Enzyme Biocatalysis, Principles and Applications. Springer Netherlands.

Ishige, T., Honda, K., Shimizu, S., **2005**. Whole organism biocatalysis. Current Opinion in Chemical Biology. 9, 174-180.

Jensen, K. F., Nygaard, P., **1975**. Purine Nucleoside Phosphorylase from Escherichia coli and Salmonella typhimurium. European Journal of Biochemistry. 51, 253-265.

Jordaan, J., Simpson, C., Brady, D., gardiner, S., Berber, I. B., Emulsion-derived particles. Vol. US 2013149730 A1. CSIR, US, **2013**.

Kaczanowski, S., Zielenkiewicz, P., **2010**. Why similar protein sequences encode similar three-dimensional structures? Theoretical Chemistry Accounts. 125, 643-650.

Kaminski, P. A., **2002**. Functional cloning, heterologous expression, and purification of two different N-deoxyribosyltransferases from Lactobacillus helveticus. Journal of Biological Chemistry. 277, 14400-14407.

Kaminski, P. A., Dacher, P., Dugué, L., Pochet, S., **2008**. In Vivo Reshaping the Catalytic Site of Nucleoside 2'-Deoxyribosyltransferase for Dideoxy- and Didehydronucleosides via a Single Amino Acid Substitution. Journal of Biological Chemistry. 283, 20053-20059.

Kang, Y. N., Zhang, Y., Allan, P. W., Parker, W. B., Ting, J. W., Chang, C. Y., Ealick, S. E., **2010**. Structure of grouper iridovirus purine nucleoside phosphorylase. Acta Crystallogr D Biol Crystallogr. 66, 155-62.

Katchalski-Katzir, E., Kraemer, D. M., **2000**. Eupergit® C, a carrier for immobilization of enzymes of industrial potential. Journal of Molecular Catalysis B: Enzymatic. 10, 157-176.

Kirk, K. L., **2006**. Fluorine in medicinal chemistry: Recent therapeutic applications of fluorinated small molecules. Journal of Fluorine Chemistry. 127, 1013-1029.

Komatsu, H., Araki, T., **2005**. Efficient chemo-enzymatic syntheses of pharmaceutically useful unnatural 2'-deoxynucleosides. Nucleosides, Nucleotides and Nucleic Acids. 24, 1127-1130.

Konstantinova, I. D., Antonov, K. V., Fateev, I. V., Miroshnikov, A. I., Stepchenko, V. A., Baranovsky, A. V., Mikhailopulo, I. A., **2011**. A Chemo-Enzymatic Synthesis of β-d-Arabinofuranosyl Purine Nucleosides. Synthesis. 2011, 1555-1560.

Konstantinova, I. D., Leont'eva, N. A., Galegov, G. A., Ryzhova, O. I., Chuvikovskii, D. V., Antonov, K. V., Esipov, R. S., Taran, S. A., Verevkina, K. N., Feofanov, S. A., Miroshnikov, A. I., **2004**. Ribavirin: Biotechnological Synthesis and Effect on the Reproduction of Vaccinia Virus. Russian Journal of Bioorganic Chemistry. 30, 553-560.

Krause, M., Ukkonen, K., Haataja, T., Ruottinen, M., Glumoff, T., Neubauer, A., Neubauer, P., Vasala, A., **2010**. A novel fed-batch based cultivation method provides high cell-density and improves yield of soluble recombinant proteins in shaken cultures. Microbial Cell Factories. 9.

Leslie, C., Eskin, E., Noble, W. S., **2002**. The spectrum kernel: a string kernel for SVM protein classification. Pac Symp Biocomput. 564-75.

Lewkowicz, E. S., Martínez, N., Rogert, M. C., Porro, S., Iribarren, A. M., **2000**. An improved microbial synthesis of purine nucleosides. Biotechnol Lett. 22, 1277-1280.

Li, N., Smith, T. J., Zong, M.-H., **2010**. Biocatalytic transformation of nucleoside derivatives. Biotechnology Advances. 28, 348-366.

Liu, P., Sharon, A., Chu, C. K., **2008**. Fluorinated Nucleosides: Synthesis and Biological Implication. J Fluor Chem. 129, 743-766.

Manoharan, M., **1999**. 2'-Carbohydrate modifications in antisense oligonucleotide therapy: importance of conformation, configuration and conjugation. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression. 1489, 117-130.

Mateo, C., Abian, O., Fernández-Lorente, G., Pedroche, J., Fernández-Lafuente, R., Guisan, J. M., **2002**. Epoxy Sepabeads: A Novel Epoxy Support for Stabilization of Industrial Enzymes via Very Intense Multipoint Covalent Attachment. Biotechnology Progress. 18, 629-634.

Mateo, C., Grazu, V., Pessela, B. C., Montes, T., Palomo, J. M., Torres, R., Lopez-Gallego, F., Fernandez-Lafuente, R., Guisan, J. M., **2007a**. Advances in the design of new epoxy supports for enzyme immobilization-stabilization. Biochem Soc Trans. 35, 1593-601.

Mateo, C., Palomo, J. M., Fernandez-Lorente, G., Guisan, J. M., Fernandez-Lafuente, R., **2007b**. Improvement of enzyme activity, stability and selectivity via immobilization techniques. Enzyme and Microbial Technology. 40, 1451-1463.

Mendieta, J., Martín-Santamaría, S., Priego, E.-M., Balzarini, J., Camarasa, M.-J., Pérez-Pérez, M.-J., Gago, F., **2004**. Role of Histidine-85 in the Catalytic Mechanism of Thymidine Phosphorylase As Assessed by Targeted Molecular Dynamics Simulations and Quantum Mechanical Calculations†. Biochemistry. 43, 405-414.

Merino, P., **2013**. Chemical Synthesis of Nucleoside Analogues. Wiley.

Mikhailopulo, I. A., **2007**. Biotechnology of Nucleic Acid Constituents - State of the Art and Perspectives. Current Organic Chemistry. 11, 317-335.

Mikhailopulo, I. A., Miroshnikov, A. I., **2010**. New trends in nucleoside biotechnology. Acta Naturae. 2, 36-59.

Mikhailopulo, I. A., Miroshnikov, A. I., **2011**. Biologically important nucloesides: modern trends in biotechnology and application. Mendeleev Communications. 21, 57-68.

Mikhailopulo, I. A., Miroshnikov, A. I., Some recent findings in the biotechnology of biologically important nucleosides. In: S. Komisarenko, (Ed.), Biochemistry and Biotechnology for Modern Medicine. Publishing House Moskalenko O.M., Kiev, Ukraine, **2013**, pp. 328-353.

Montilla Arevalo, R., Deroncelé Thomas, V. M., López Gómez, C., Pascual Gilabert, M., Estévez Company, C., Castells Boliart, J., Thermostable biocatalyst combination for nucleoside synthesis. In: E. P. Office, (Ed.), Vol. EP 2338985 A1. Institut Univ. de Ciència i Tecnologia, European, **2011**.

Mushegian, A. R., Koonin, E. V., **1994**. Unexpected sequence similarity between nucleosidases and phosphoribosyltransferases of different specificity. Protein Science. 3, 1081-1088.

Nóbile, M., Médici, R., Terreni, M., Lewkowicz, E. S., Iribarren, A. M., **2012**. Use of Citrobacter koseri whole cells for the production of arabinonucleosides: A larger scale approach. Process Biochemistry. 47, 2182-2188.

Nazina, T. N., Tourova, T. P., Poltaraus, A. B., Novikova, E. V., Grigoryan, A. A., Ivanova, A. E., Lysenko, A. M., Petrunyaka, V. V., Osipov, G. A., Belyaev, S. S., Ivanov, M. V., **2001**. Taxonomic study of aerobic thermophilic bacilli: descriptions of Geobacillus subterraneus gen. nov., sp nov and Geobacillus uzenensis sp nov from petroleum reservoirs and transfer of Bacillus stearothermophilus Bacillus thermocatenulatus, Bacillus thermoleovorans, Bacillus kaustophilus, Bacillus thermoglucosidasius and Bacillus thermodenitrificans to Geobacillus as the new combinations G-stearothermophilus, G-thermocatenulatus, G-thermoleovorans, G-kaustophilus, G-

thermoglucosidasius and G-thermodenitrificans. International Journal of Systematic and Evolutionary Microbiology. 51, 433-446.

Nguyen, V. D., Hatahet, F., Salo, K. E., Enlund, E., Zhang, C., Ruddock, L. W., **2011**. Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of E.coli. Microbial Cell Factories. 10, 1.

Oshima, T., Imahori, K., **1974**. Description of Thermus-Thermophilus (Yoshida and Oshima) Comb-Nov, a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. International Journal of Systematic Bacteriology. 24, 102-112.

Ouwerkerk, N., Steenweg, M., de Ruijter, M., Brouwer, J., van Boom, J. H., Lugtenburg, J., Raap, J., **2002**. One-Pot Two-Step Enzymatic Coupling of Pyrimidine Bases to 2-Deoxy-d-ribose-5-phosphate. A New Strategy in the Synthesis of Stable Isotope Labeled Deoxynucleosides. The Journal of Organic Chemistry. 67, 1480-1489.

Owusu, R. K., Cowan, D. A., **1989**. Correlation between microbial protein thermostability and resistance to denaturation in aqueous: organic solvent two-phase systems. Enzyme and Microbial Technology. 11, 568-574.

Pace, N. R., **2006**. Time for a change. Nature. 441, 289-289.

Prakash, T. P., Bhat, B., **2007**. 2'-Modified oligonucleotides for antisense therapeutics. Current Topics in Medicinal Chemistry. 7, 641-649.

Pugmire, M. J., Cook, W. J., Jasanoff, A., Walter, M. R., Ealick, S. E., **1998**. Structural and theoretical studies suggest domain movement produces an active conformation of thymidine phosphorylase. Journal of Molecular Biology. 281, 285-299.

Pugmire, M. J., Ealick, S. E., **1998**. The crystal structure of pyrimidine nucleoside phosphorylase in a closed conformation. Structure. 6, 1467-79.

Pugmire, M. J., Ealick, S. E., **2002**. Structural analyses reveal two distinct families of nucleoside phosphorylases. Biochemical Journal. 361, 1-25.

Rivero, C. W., De Benedetti, E. C., Sambeth, J. E., Lozano, M. E., Trelles, J. A., **2012**. Biosynthesis of anti-HCV compounds using thermophilic microorganisms. Bioorganic & Medicinal Chemistry Letters. 22, 6059-6062.

Robak, T., Korycka, A., Lech-Maranda, E., Robak, P., **2009**. Current Status of Older and New Purine Nucleoside Analogues in the Treatment of Lymphoproliferative Diseases. Molecules. 14, 1183-1226.

Robak, T., Lech-Maranda, E., Korycka, A., Robak, E., **2006**. Purine nucleoside analogs as immunosuppressive and antineoplastic agents: mechanism of action and clinical activity. Current Medicinal Chemistry. 13, 3165-3189.

Rocchietti, S., Ubiali, D., Terreni, M., Albertini, A. M., Fernandez-Lafuente, R., Guisan, J. M., Pregnolato, M., **2004**. Immobilization and stabilization of recombinant multimeric uridine and purine nucleoside phosphorylases from Bacillus subtilis. Biomacromolecules. 5, 2195-200.

Roshevskaia, L. A., Barai, V. N., Zinchenko, A. I., Kvasiuk, E. I., Mikhailopulo, I. A., **1986**. Preparative synthesis of the antiviral nucleoside 9-beta-D-arabinofuranosyladenine by using bacterial cells. Antibiot Med Biotekhnol. 31, 174-8.

Sako, Y., Nomura, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T., Maruyama, T., **1996**. Aeropyrum pernix gen nov, sp nov, a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 degrees C. International Journal of Systematic Bacteriology. 46, 1070-1077.

Saunders, P. P., Wilson, B. A., Saunders, G. F., **1969**. Purification and Comparative Properties of a Pyrimidine Nucleoside Phosphorylase from Bacillus Stearothermophilus. Journal of Biological Chemistry. 244, 3691-3697.

Seela, F., Xiong, H., Leonard, P., Budow, S., **2009**. 8-Aza-7-deazaguanine nucleosides and oligonucleotides with octadiynyl side chains: synthesis, functionalization by the azide-alkyne 'click' reaction and nucleobase specific fluorescence quenching of coumarin dye conjugates. Organic & Biomolecular Chemistry. 7, 1374-1387.

Serra, I., Bavaro, T., Cecchini, D. A., Daly, S., Albertini, A. M., Terreni, M., Ubiali, D., **2013a**. A comparison between immobilized pyrimidine nucleoside phosphorylase from Bacillus subtilis and thymidine phosphorylase from Escherichia coli in the synthesis of 5-substituted pyrimidine 2'-deoxyribonucleosides. Journal of Molecular Catalysis B: Enzymatic. 95, 16-22.

Serra, I., Serra, C. D., Rocchietti, S., Ubiali, D., Terreni, M., **2011**. Stabilization of thymidine phosphorylase from Escherichia coli by immobilization and post immobilization techniques. Enzyme and Microbial Technology. 49, 52-58.

Serra, I., Ubiali, D., Piškur, J., Christoffersen, S., Lewkowicz, E. S., Iribarren, A. M., Albertini, A. M., Terreni, M., **2013b**. Developing a Collection of Immobilized Nucleoside Phosphorylases for the Preparation of Nucleoside Analogues: Enzymatic Synthesis of Arabinosyladenine and 2',3'-Dideoxyinosine. ChemPlusChem. 78, 157-165.

Sgarrella, F., Frassetto, L., Allegnini, S., Camici, M., Carta, M. C., Fadda, P., Tozzi, M. G., Ipata, P. L., **2007**. Characterization of the adenine nucleoside specific phosphorylase of Bacillus cereus. Biochimica Et Biophysica Acta-General Subjects. 1770, 1498-1505.

Sharp, P. M., Li, W. H., **1987**. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281-95.

Sinisterra, J.-V., Alcántara, A.-R., Almendros, M., Hernáiz, M. J., Sánchez-Montero, J.-M., Trelles, J., Flickinger, M. C., Enzyme-Catalyzed Synthesis of Nonnatural or Modified Nucleosides. Encyclopedia of Industrial Biotechnology. John Wiley & Sons, Inc., **2009**.

Siurkus, J., Panula-Perala, J., Horn, U., Kraft, M., Rimseliene, R., Neubauer, P., **2010**. Novel approach of high cell density recombinant bioprocess development: Optimisation and scale-up from microlitre to pilot scales while maintaining the fed-batch cultivation mode of E. coli cultures. Microbial Cell Factories. 9, 35.

Sivets, G. G., Kalinichenko, E. N., Mikhailopulo, I. A., **2006**. Synthesis of C2 '-beta-fluoro-substituted adenine nucleosides via pivaloyl derivatives of adenosine and 3'-deoxyadenosine. Letters in Organic Chemistry. 3, 402-408.

Souleimanian, N., Deleavey, G. F., Soifer, H., Wang, S., Tiemann, K., Damha, M. J., Stein, C. A., **2012**. Antisense 2[prime]-Deoxy, 2[prime]-Fluroarabino Nucleic Acids (2[prime]F-ANAs) Oligonucleotides: In Vitro Gymnotic Silencers of Gene Expression Whose Potency Is Enhanced by Fatty Acids. Mol Ther Nucleic Acids. 1, e43.

Stepchenko, V. A., Seela, F., Esipov, R. S., Miroshnikov, A. I., Sokolov, Y. A., Mikhailopulo, I. A., **2012**. Enzymatic Synthesis of 2'-Deoxy-β-d-ribonucleosides of 8-Azapurines and 8-Aza-7-deazapurines. Synlett. 23, 1541-1545.

Stoeckler, J. D., Cambor, C., Parks, R. E., **1980**. Human erythrocytic purine nucleoside phosphorylase: reaction with sugar-modified nucleoside substrates. Biochemistry. 19, 102-107.

Stoeckler, J. D., Poirot, A. F., Smith, R. M., Parks, R. E., Jr., Ealick, S. E., Takabayashi, K., Erion, M. D., **1997**. Purine nucleoside phosphorylase. 3. Reversal of purine base specificity by site-directed mutagenesis. Biochemistry. 36, 11749-56.

Suzuki, Y., Kishigami, T., Inoue, K., Mizoguchi, Y., Eto, N., Takagi, M., Abe, S., **1983**. Bacillus thermoglucosidasius sp. nov., a New Species of Obligately Thermophilic Bacilli. Systematic and Applied Microbiology. 4, 487-495.

Szeker, K., Nucleoside phosphorylases from thermophiles - Recombinant expression and biocatalytic use for modified nucleosides. Institute of Biotechnology, Vol. Doctor. Technische Universität Berlin, Berlin, **2012**.

Szeker, K., Niemitalo, O., Casteleijn, M. G., Juffer, A. H., Neubauer, P., **2011**. High-temperature cultivation and 5' mRNA optimization are key factors for the efficient

overexpression of thermostable Deinococcus geothermalis purine nucleoside phosphorylase in Escherichia coli. J Biotechnol. 156, 268-74.

Szeker, K., Zhou, X., Schwab, T., Casanueva, A., Cowan, D., Mikhailopulo, I. A., Neubauer, P., **2012**. Comparative investigations on thermostable pyrimidine nucleoside phosphorylases from *Geobacillus thermoglucosidasius* and *Thermus thermophilus*. Journal of Molecular Catalysis B: Enzymatic. 84, 27-34.

Taran, S., Verevkina, K., Feofanov, S., Miroshnikov, A., **2009**. Enzymatic transglycosylation of natural and modified nucleosides by immobilized thermostable nucleoside phosphorylases from *Geobacillus stearothermophilus*. Russian Journal of Bioorganic Chemistry. 35, 739-745.

Taverna-Porro, M., Bouvier, L. A., Pereira, C. A., Montserrat, J. M., Iribarren, A. M., **2008**. Chemoenzymatic preparation of nucleosides from furanoses. Tetrahedron Letters. 49, 2642-2645.

Tennilä, T., Azhayeva, E., Vepsäläinen, J., Laatikainen, R., Azhayev, A., Mikhailopulo, I. A., **2000**. Oligonucleotides Containing 9-(2-Deoxy-2-Fluoro-β-D-Arabinofuranosyl)-Adenine and -Guanine: Synthesis, Hybridization and Antisense Properties. Nucleosides, Nucleotides and Nucleic Acids. 19, 1861-1884.

Thomas, H. J., Tiwari, K. N., Clayton, S. J., Secrist, J. A., Montgomery, J. A., **1994**. Synthesis and Biologic Activity of Purine 2'-Deoxy-2'-fluoro-ribonucleosides. Nucleosides & Nucleotides. 3, 309-323.

Tran, T. H., Christoffersen, S., Allan, P. W., Parker, W. B., Piskur, J., Serra, I., Terreni, M., Ealick, S. E., **2011**. The crystal structure of Streptococcus pyogenes uridine phosphorylase reveals a distinct subfamily of nucleoside phosphorylases. Biochemistry. 50, 6549-58.

Trelles, J. A., Valino, A. L., Runza, V., Lewkowicz, E. S., Iribarren, A. M., **2005**. Screening of catalytically active microorganisms for the synthesis of 6-modified purine nucleosides. Biotechnol Lett. 27, 759-63.

Turkman, N., Gelovani, J. G., Alauddin, M. M., **2010**. A novel method for stereospecific fluorination at the 2 '-arabino-position of pyrimidine nucleoside: synthesis of [(18)F]-FMAU. Journal of Labelled Compounds & Radiopharmaceuticals. 53, 782-786.

Tuttle, J. V., Tisdale, M., Krenitsky, T. A., **1993**. Purine 2'-deoxy-2'-fluororibosides as antiinfluenza virus agents. J Med Chem. 36, 119-125.

Twala, B. V., Sewell, B. T., Jordaan, J., **2012**. Immobilisation and characterisation of biocatalytic co-factor recycling enzymes, glucose dehydrogenase and NADH oxidase, on aldehyde functional ReSyn™ polymer microspheres. Enzyme and Microbial Technology. 50, 331-336.

Ubiali, D., Rocchietti, S., Scaramozzino, F., Terreni, M., Albertini, A. M., Fernández-Lafuente, R., Guisán, J. M., Pregnolato, M., **2004**. Synthesis of 2'-Deoxynucleosides by Transglycosylation with New Immobilized and Stabilized Uridine Phosphorylase and Purine Nucleoside Phosphorylase. Advanced Synthesis & Catalysis. 346, 1361-1366.

Ubiali, D., Serra, C. D., Serra, I., Morelli, C. F., Terreni, M., Albertini, A. M., Manitto, P., Speranza, G., **2012**. Production, Characterization and Synthetic Application of a Purine Nucleoside Phosphorylase from Aeromonas hydrophila. Advanced Synthesis & Catalysis. 354, 96-104.

Ukkonen, K., Vasala, A., Ojamo, H., Neubauer, P., **2011**. High-yield production of biologically active recombinant protein in shake flask culture by combination of enzyme-based glucose delivery and increased oxygen transfer. Microbial Cell Factories. 10, 107.

Utagawa, T., Morisawa, H., Yamanaka, S., Yamazaki, A., Yoshinaga, F., Hirose, Y., **1985**. Properties of Nucleoside Phosphorylase from Enterobacter aerogenes. Agricultural and Biological Chemistry. 49, 3239-3246.

Vagenende, V., Yap, M. G., Trout, B. L., **2009**. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry. 48, 11084-96.

Vieille, C., Zeikus, G. J., **2001**. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. Microbiology and Molecular Biology Reviews. 65, 1-43.

Vorbrüggen, H., Ruh-Pohlenz, C., Synthesis of Nucleosides. Organic Reactions. John Wiley & Sons, Inc., **2004**.

Watts, J. K., Corey, D. R., **2012**. Silencing disease genes in the laboratory and the clinic. The Journal of Pathology. 226, 365-379.

Watts, J. K., Damha, M. J., **2008**. 2 ' F-arabinonucleic acids (2 ' F-ANA) - History, properties, and new frontiers. Canadian Journal of Chemistry-Revue Canadienne De Chimie. 86, 641-656.

Wei, X., Ding, Q., Ou, L., Zhang, L., Wang, C., **2008a**. Two-step Enzymatic Synthesis of Guanine Arabinoside. Chinese Journal of Chemical Engineering. 16, 934-937.

Wei, X. K., Ding, Q. B., Zhang, L., Guo, Y. L., Ou, L., Wang, C. L., **2008b**. Induction of nucleoside phosphorylase in Enterobacter aerogenes and enzymatic synthesis of adenine arabinoside. J Zhejiang Univ Sci B. 9, 520-6.

Yamada, K., Matsumoto, N., Hayakawa, H., **2009**. Stereoselective Synthesis of 2-Deoxy-2-Fluoroarabinofuranosyl-a-1-Phosphate and Its Application to the Synthesis of 2'-Deoxy-2'-Fluoroarabinofuranosyl Purine Nucleosides by a Chemo-Enzymatic Method. Nucleosides, Nucleotides and Nucleic Acids. 28, 1117-1130.

Yamazaki, S., Yamazaki, J., Nishijima, K., Otsuka, R., Mise, M., Ishikawa, H., Sasaki, K., Tago, S., Isono, K., **2006**. Proteome analysis of an aerobic hyperthermophilic crenarchaeon, Aeropyrum pernix K1. Mol Cell Proteomics. 5, 811-23.

Yon, J., Perahia, D., Ghelis, C., **1998**. Conformational dynamics and enzyme activity. Biochimie. 80, 33-42.

Yuan, Z., Bailey, T. L., Teasdale, R. D., **2005**. Prediction of protein B-factor profiles. Proteins: Structure, Function, and Bioinformatics. 58, 905-912.

Zaitseva, G. V., Zinchenko, A. I., Barai, V. N., Pavlova, N. I., Boreko, E. I., Mikhailopulo, I. A., **1999**. Chemical and enzymatic synthesis and antiviral properties of 2'-deoxy-2'-fluoroguanosine. Nucleosides Nucleotides. 18, 687-8.

Zhao, Y., Caspers, M. P., Abee, T., Siezen, R. J., Kort, R., **2012**. Complete Genome Sequence of Geobacillus thermoglucosidans TNO-09.020, a Thermophilic Sporeformer Associated with a Dairy-Processing Environment. Journal of Bacteriology. 194, 4118.

Zhou, X., Szeker, K., Janocha, B., Böhme, T., Albrecht, D., Mikhailopulo, I. A., Neubauer, P., **2013**. Recombinant purine nucleoside phosphorylases from thermophiles: preparation, properties and activity towards purine and pyrimidine nucleosides. FEBS Journal. 280, 1475-1490.

Zhu, S., Ren, L., Wang, J., Zheng, G., Tang, P., **2012a**. Two-step efficient synthesis of 5-methyluridine via two thermostable nucleoside phosphorylase from Aeropyrum pernix. Bioorganic & Medicinal Chemistry Letters. 22, 2102-2104.

Zhu, S., Song, D., Gong, C., Tang, P., Li, X., Wang, J., Zheng, G., **2012b**. Biosynthesis of nucleoside analogues via thermostable nucleoside phosphorylase. Applied Microbiology and Biotechnology. 1-10.

Zuffi, G., Ghisotti, D., Oliva, I., Capra, E., Frascotti, G., Tonon, G., Orsini, G., **2004**. Immobilized biocatalysts for the production of nucleosides and nucleoside analogues by enzymatic transglycosylation reactions. Biocatalysis and Biotransformation. 22, 25-33.

Paper I

Journal of Molecular Catalysis B: Enzymatic 84 (2012) 27–34

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

Comparative investigations on thermostable pyrimidine nucleoside phosphorylases from *Geobacillus thermoglucosidasius* and *Thermus thermophilus*

Kathleen Szeker[a,*], Xinrui Zhou[a], Thomas Schwab[b], Ana Casanueva[c], Don Cowan[c], Igor A. Mikhailopulo[d], Peter Neubauer[a]

[a] Department of Bioprocess Engineering, Institute of Biotechnology, Technische Universität Berlin, Ackerstr. 71-76, ACK24, 13355 Berlin, Germany
[b] Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
[c] Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Private Bag X17, Bellville 7535, Cape Town, South Africa
[d] Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Acad. Kuprevicha 5/2, 220141 Minsk, Belarus

ARTICLE INFO

Article history:
Available online 25 February 2012

Keywords:
Biocatalysis
Thermostable pyrimidine nucleoside phosphorylase
Pyrimidine nucleoside
2′-Deoxyfluoro nucleosides
Thermus thermophilus
Geobacillus thermoglucosidasius

ABSTRACT

The recombinant expression and biocatalytic characterization of two thermostable pyrimidine nucleoside phosphorylases (PyNP), isolated from *Geobacillus thermoglucosidasius* (Gt) and *Thermus thermophilus* (Tt) is described. Both enzymes are highly thermostable (half life of GtPyNP is 1.6 h at 70 °C, half life of TtPyNP is >24 h at 80 °C). Kinetic parameters for the phosphorolysis of natural substrates were determined for GtPyNP at 60 °C (K_m for uridine 2.3 mM, K_m for thymidine 1.3 mM) and TtPyNP at 80 °C (K_m for uridine 0.15 mM, K_m for thymidine 0.43 mM). The k_{cat} values for uridine are almost identical for both enzymes (ca. 277 s^{-1}), while the k_{cat} value for thymidine is about 8 times higher for TtPyNP than for GtPyNP (679 s^{-1} vs. 83 s^{-1}).

Both enzymes were tested towards the ability to catalyze the phosphorolytic cleavage of 2′-fluorosubstituted pyrimidine nucleosides – a prerequisite for the efficient synthesis of a number of relevant purine nucleoside analogues. GtPyNP showed poor activity towards 2′-deoxy-2′-fluorouridine ($dUrd_{2'F}$; 0.4% substrate conversion after 30 min), and the phosphorolysis of the epimeric counterpart 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)uracil ($dUrd^{2'F}$) could not be detected at all. By contrast, TtPyNP showed dramatically higher conversion rates (15.6% and 1.6% conversion in 30 min of both substrates, respectively). The amount of converted pyrimidine nucleosides increased significantly with time. After 17 h 65% of $dUrd_{2'F}$ and 46% of $dUrd^{2'F}$ was phosphorolytically cleaved.

Our results demonstrate the potential of TtPyNP as a biocatalyst in transglycosylation reactions aiming at the production of 2′-fluorosubstituted purine nucleosides that are highly bioactive but hardly accessible by chemical methods.

1. Introduction

Pyrimidine nucleoside phosphorylases (PyNP; EC 2.4.2.2) are homodimeric enzymes that are involved in the nucleotide salvage pathway of some lower organisms [1–3]. In the presence of phosphate ions, PyNP catalyzes the reversible phosphorolytic cleavage of the glycosidic bond of pyrimidine nucleosides or closely related derivatives thereof according to the following general reaction scheme (Scheme 1).

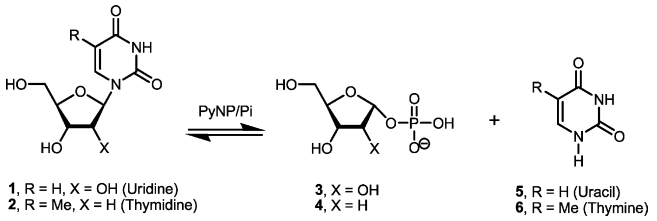
Structurally, PyNP is closely related to human thymidine phosphorylases (TP; EC 2.4.2.4) – an angiogenic enzyme that was found to be identical to the platelet-derived endothelial cell growth factor in humans [4,5]. Together, TP and PyNP form the nucleoside phosphorylase-II family, which share a common two-domain subunit fold and a high level of sequence identity [3]. Despite the similarity of the reaction catalyzed, uridine phosphorylase (UP; EC 2.4.2.3) belongs to the phosphorylase-I family with distinct structural characteristics. From the catalytical point of view, TP is distinguished from UP due to its high specificity for the 2′-deoxy-D-ribofuranose moiety of pyrimidine nucleosides [3]. By contrast, PyNP does not discriminate between uridine (**1**) and thymidine (**2**) and accepts both compounds as natural substrates [6].

This rather broad substrate specificity makes PyNP a versatile biocatalyst suitable for certain synthetic applications, e.g. for the

Abbreviations: PyNP, pyrimidine nucleoside phosphorylase; PNP, purine nucleoside phosphorylase; NP, nucleoside phosphorylase; TP, thymidine phosphorylase; UP, uridine phosphorylase; Tt, *Thermus thermophilus*; Gt, *Geobacillus thermoglucosidasius*; $dUrd_{2'F}$, 2′-deoxy-2′-fluorouridine; $dUrd^{2'F}$, 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)uracil.

* Corresponding author. Tel.: +49 30 314 72521; fax: +49 30 314 27577.
E-mail addresses: k.szeker@mailbox.tu-berlin.de (K. Szeker), xinrui.zhou@gmail.com (X. Zhou), thomas.schwab@biologie.uni-regensburg.de (T. Schwab), acasanueva@uwc.ac.za (A. Casanueva), dcowan@uwc.ac.za (D. Cowan), igormikh@iboch.bas-net.by (I.A. Mikhailopulo), peter.neubauer@tu-berlin.de (P. Neubauer).

1381-1177/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2012.02.006



Scheme 1.

enzymatic synthesis of nucleosides. Of particular interest is the synthesis of modified nucleosides, that can be used as pharmaceutical agents for the treatment of viral infections and cancer, as well as in molecular biological techniques and diagnostics [7–10]. For their synthesis, efficient chemo-enzymatic approaches have been developed, among them the enzymatic transglycosylation of chemically modified nucleosides [11]. Scheme 2 illustrates the enzymatic synthesis of modified purine nucleosides as an example: by the concerted action of a PyNP and a purine nucleoside phosphorylase (PNP; EC 2.4.2.1), a pentofuranosyl moiety is transferred from a pyrimidine nucleoside that serves as pentofuranosyl donor to a heterocyclic purine base that serves as pentofuranosyl acceptor. It is obvious that the lack of enzymes with broad substrate specificities is limiting the application spectrum of this strategy. In addition, the efficiency of such processes, when run at high temperatures, is often reduced due to the thermal lability of the biocatalysts.

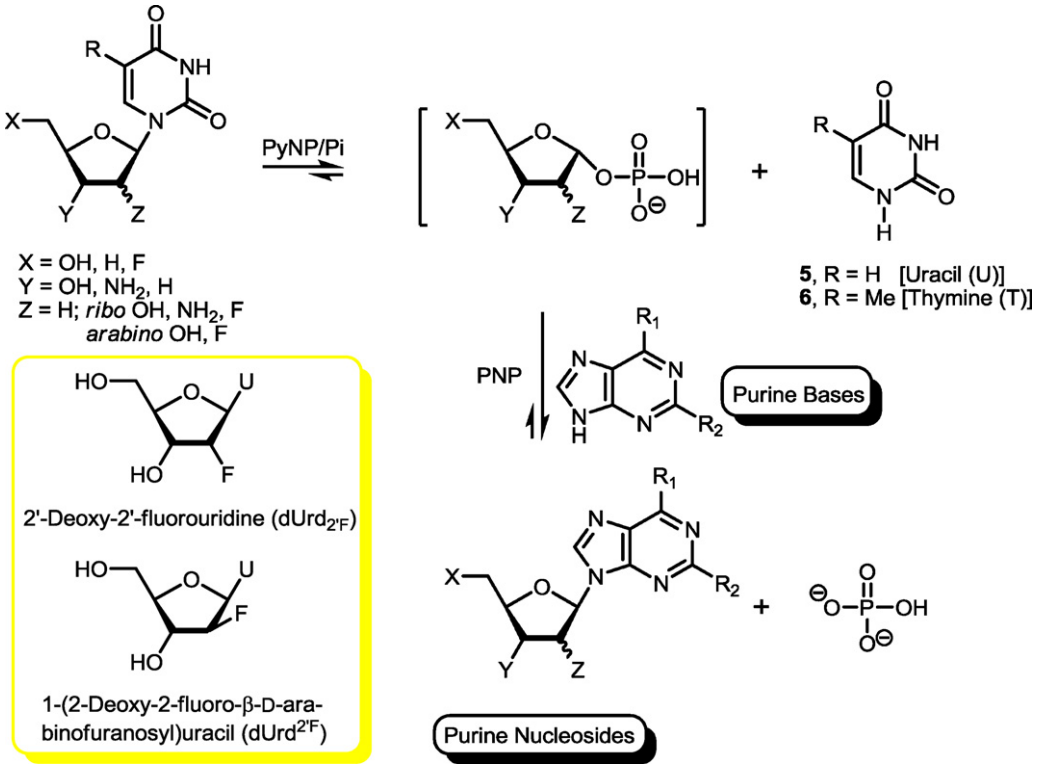
The search for enzymes with improved properties is hence of prime importance in this field of research and prompted us to investigate biocatalytical characteristics of novel nucleoside phosphorylases. In contrast to PNPs that have been extensively studied from a multitude of mesophilic and thermophilic (micro) organisms [12–15], PyNPs have been hardly studied in detail. Only few examples can be found in the scientific literature: PyNP from *Bacillus subtilis* [16], *Geobacillus stearothermophilus* [1,2,6] and from *T. thermophilus* [17]. With respect to practical use as biocatalyst, only the *G. stearothermophilus* enzyme has been described [18].

In the present study, we report on the recombinant expression and characterization of two additional thermostable PyNPs, derived from *G. thermoglucosidasius* 11955 (GtPyNP) and *T. thermophilus* HB27 (TtPyNP), respectively. Albeit the expression and crystallization of PyNP from *T. thermophilus* HB8 has been reported [17], information about the biocatalytic characterization is not available. Here, we investigate thermal characteristics and kinetic parameters as well as the ability to phosphorolyze 2′-fluorosubstituted pyrimidine nucleosides, *viz.*, 2′-deoxy-2′-fluorouridine ($dUrd_{2'F}$) and 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)uracil ($dUrd^{2'F}$) (Scheme 2), which are of special interest, as sugar donors, for purine nucleoside synthesis. Our data reveal striking differences of the conversion rates depending on the substrate analogue and the biocatalyst applied.

2. Experimental

2.1. Molecular biology

G. thermoglucosidasius 11955 was grown for 1.5 days at 52 °C in Luria Broth (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, pH 7.0), whereas *T. thermophilus* HB27 was grown as previously described [19]. Genomic DNA of both strains was isolated using standard protocols [20]. The genes coding for PyNP in *G. thermoglucosidasius* (GenBank accession no. ZP_06809030) and *T. thermophilus* (GenBank accession number AAS81754.1) were amplified using Pfu DNA polymerase (Fermentas, Lithuania). The following primer pair was used for the isolation of the PyNP gene from the *G. thermoglucosidasius* genome: 5′ ACTAGGGATCCATGGTCGATTTAATTGCGA 3′ (*Bam*HI site underlined) and 5′ AGCATGCGGCCGCTTATGAAATGGTTTCGTATATA 3′ (*Not*I site underlined). The primer pair used for the *T. thermophilus* gene was: 5′ ACTAGGGATCCAACCCCGTGGTCTTCATC 3′ (*Bam*HI



Scheme 2.

site underlined) and 5′ AGCATGCGGCCGCCTAGATGGCCTCCAGGA 3′ (*Not*I site underlined). The PyNP encoding genes were cloned via *Bam*HI/*Not*I digestion (FastDigest restriction endonucleases, Fermentas, Lithuania) and subsequent ligation (T4 DNA Ligase, Roche) into a derivative of the expression vector pCTUT7 [21]. This vector is characterized by an IPTG inducible *lac* promoter derivative and a pBR322 origin of replication. We have modified this vector before by substituting the chloramphenicol resistance cassette by the plasmid stabilizing *parB* locus [22] and an ampicillin resistance cassette. In addition we introduced a sequence encoding a hexahistidine tag connected to the 5′-end of target gene's transcript. The resulting expression vectors were transformed into the *Escherichia coli* BL21 strain (Novagen).

2.2. Bioinformatics and homology modeling

Amino acid sequence identities were assessed with the protein basic local alignment tool (BLAST) [23] of the NCBI web server. For multiple sequence alignments the ClustalW2 program from the EBI web server was used [24]. Three-dimensional protein structures of the target proteins were built by homology modeling using the Swiss-model workspace [25]. These 3D models were superposed and visualized with the CCP 4 mg molecular-graphics software [26].

2.3. Overexpression in E. coli

Recombinant *E. coli* strains were cultivated in terrific broth (TB) medium [27] or EnPresso® medium (BioSilta, Finland) using Ultra Yield Flasks™ and AirOTop™ seals (Thomson Instrument Company, USA). Recombinant protein expression was induced by addition of IPTG to a final concentration of 100 μM, after 2.5 h cultivation time (TB medium) or after overnight cultivation (EnPresso medium). Cells were harvested (centrifugation at 16,000 × *g*, 5 min, 4 °C) 3 h after induction (TB medium) or 24 h after induction (EnPresso medium) and stored at −20 °C.

2.4. Cell disruption and purification

Cell pellets were re-suspended in NPI-10 binding buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) at 5 ml/g wet weight. This cell suspension was sonicated on ice using a UP200S sonicator (Hielscher Ultrasonics GmbH). The sonication was performed twice with 30% power input for 3 min in 30 s intervals using a sonotrode of 7 mm in diameter.

Cell lysates were centrifuged (20,000 × *g*, 15 min, 4 °C) to separate soluble from insoluble fraction. The soluble portion of the cell lysate was heated for 15 min at 60 °C (GtPyNP) or 80 °C (TtPyNP). Coagulated proteins were removed by centrifugation (20,000 × *g*, 15 min, 4 °C). The target proteins were further purified by immobilized metal ion affinity chromatography using a 5 ml Ni-NTA Superflow cartridge (Qiagen), subsequently the excess of salt and imidazole was removed by the use of a HiPrep 26/10 Desalting column (GE Healthcare). The purified protein solution was aliquoted, rapidly frozen in liquid nitrogen and stored at −80 °C.

2.5. Protein analytics

The protein concentration of the purified protein solution was determined by measuring the absorption at 280 nm (Nanodrop, Thermo Scientific). The absorption coefficients were theoretically calculated from the amino acid sequence (Vector NTI software, Invitrogen) and were 0.42 AU (GtPyNP) and 0.56 AU (TtPyNP) for a solution of 1 mg ml^{-1}.

The molecular weight marker for SDS-PAGE analysis was purchased from Fermentas (Lithuania).

Protein unfolding: Thermal denaturation of 10 μM purified TtPyNP protein dissolved in potassium phosphate (50 mM, pH 7.5) was monitored with a Jasco J-815 circular dichroism (CD) spectrometer in a 0.1 cm cuvette by following the loss of ellipticity at 220 nm. Unfolding was induced by raising the temperature in 0.1 °C increments at a ramp rate of 1 °C min^{-1} with a Peltier-effect temperature controller. The measured ellipticity was normalized, and the apparent melting temperature (T_M^{app}) was determined. DSC experiments were performed with 43 μM TtPyNP protein dissolved in potassium phosphate (50 mM, pH 7.5) by heating the samples in a CSC 5100 Nano differential scanning calorimeter with a scan rate of 1 °C min^{-1}. The DSC data were analyzed with the program CpCalc (version 2.1, Calorimetry Sciences Corporation, 1995) to determine T_M^{app} at which half of the protein is unfolded. The irreversibility of thermal denaturation precluded thermodynamic analysis of the CD and DSC unfolding traces.

2.6. Activity assay

If not otherwise stated the activity assay was performed in 50 mM potassium phosphate buffer, pH 7.0 containing 1 mM of uridine as substrate. After 2 min of pre-heating, 1–2 μl of diluted enzyme was added per 100 μl of reaction mixture. Samples were withdrawn after defined time intervals and the reaction was immediately stopped by adding 1 volume of reaction mixture to ½ volume of 10% trichloroacetic acid. Precipitated proteins were removed by centrifugation (20,000 × *g*, 15 min) and the samples were stored at −20 °C for later analysis by HPLC.

From the HPLC results the concentration of residual pyrimidine nucleoside (substrate) and liberated pyrimidine base (product of the phosphorolytic cleavage) was retrieved. The substrate conversion was calculated with the following formula:

$$\text{Conversion}\,(\%) = \frac{\text{base}}{\text{base} + \text{nucleoside}} \times 100$$

Only conversion rates that were linear with respect to time and amount of PyNP added, were considered for further analysis. This was usually the case if not more than 10% of the substrate was converted.

2′-Deoxy-2′-fluorouridine was purchased from TCI Deutschland (Eschborn, Germany) and 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)uracil obtained from Metkinen Chemistry (Kuusisto, Finland). Other chemical reagents were purchased from Sigma–Aldrich.

2.6.1. Thermal stability

Enzyme preparations were diluted 1:100 in 50 mM potassium phosphate buffer (pH 7.0) yielding enzyme concentrations of about 46 μg ml^{-1}. These diluted enzyme solutions were aliquoted and incubated in 0.2 ml tubes in a thermocycler at the respective temperatures. After defined time intervals, tubes were withdrawn and the residual activity of the incubated enzyme solution was determined (at 60 °C for GtPyNP and at 80 °C for TtPyNP) and plotted over the reaction time. In order to determine the half life, the resulting curve was fitted (Sigma Plot) to the decay function:

$$v = v_0 \cdot e^{k_i \cdot t}$$

and subsequently the half life was calculated:

$$t_{1/2} = \ln 0.5 \cdot {k_i}^{-1}$$

2.6.2. Temperature optimum

The reaction mixture was pre-heated for 2 min at respective temperatures. The diluted enzyme solution was added and the

reaction was stopped after 3 min. 50 mM potassium phosphate buffer (pH 7.5) was used.

2.6.3. Kinetic parameters

Activity tests were performed in triplicates for at least 5 different substrate concentrations spanning 0.25–5 times of the Michaelis–Menton constant. The initial reaction rates were plotted over the substrate concentrations and fitted to the Michaelis–Menton function (Sigma Plot). From the resulting formula K_m and v_{max} were directly retrieved, and k_{cat} could be easily calculated. For details see [28].

2.6.4. Phosphorolysis of modified nucleosides

Reactions were performed in 1 mM pyrimidine nucleoside substrate solution containing 10 mM sodium phosphate buffer (pH 6.5) and an enzyme loading of ca. 0.1 mg ml^{-1} at either 60 °C or 80 °C.

2.6.5. HPLC analysis

The amount of pyrimidine nucleosides and corresponding bases was determined by following the absorption at 260 nm during HPLC analysis using a reversed phase C18 column (Gemini-Nx 5u, 150 × 4.60 mm, Phenomenex) with the following gradient: from 97% 20 mM ammonium acetate and 3% acetonitrile to 60% 20 mM ammonium acetate and 40% acetonitrile in 10 min. Under these conditions the following retention times were determined: uridine (3.2 min), uracil (2.4 min), thymidine (4.7 min), thymine (4.0 min), dUrd$^{2'F}$ (4.6 min), dUrd$_{2'F}$ (4.4 min), O^2,2′-anhydro-1-(β-D-arabinofuranosyl)uracil (2.3 min), 1-(β-D-arabinofuranosyl)uracil (3.8 min).

1-(β-D-Arabinofuranosyl)uracil was purchased from Sigma–Aldrich, O^2,2′-anhydro-1-(β-D-arabinofuranosyl)uracil was synthesized as described in [29].

3. Results and discussion

3.1. Sequence analysis and homology modeling

The proteins recombinantly expressed and characterized in this work are the pyrimidine nucleoside phosphorylases from *T. thermophilus* HB27 (GenBank accession number AAS81754.1) and from *G. thermoglucosidasius* 11955. The sequence of the latter corresponded to the PyNP sequence from *G. thermoglucosidasius* C56-YS93 (GenBank accession number ZP_06809030), with the exception of one amino acid. Since the corresponding deviation of the DNA sequence was suspected to be the result of the PCR performed for cloning, the respective codon was changed by site-directed mutagenesis to fit the data bank protein (Gln_{214} changed to Arg).

Three entries of solved crystal structures of PyNPs can be found in the protein database (www.pdb.org). Both target enzyme sequences were blasted against the amino acid sequences belonging to these pdb entries. GtPyNP aligned best with the PyNP from *G. stearothermophilus ATCC* 12980 (assigned here as GsPyNP, pdb entry 1BRW) with 78% sequence identity. TtPyNP aligned best with the PyNP from *T. thermophilus* HB8 (pdb entry 2DSJ), to which it is almost identical (approx. 98% sequence identity) but showed also a high degree of sequence identity (approx. 50%) to GsPyNP. The amino acid sequence alignment of GsPyNP, GtPyNP, TtPyNP and *E. coli* thymidine phosphorylase (EcTP) is shown in Fig. 1. Amino

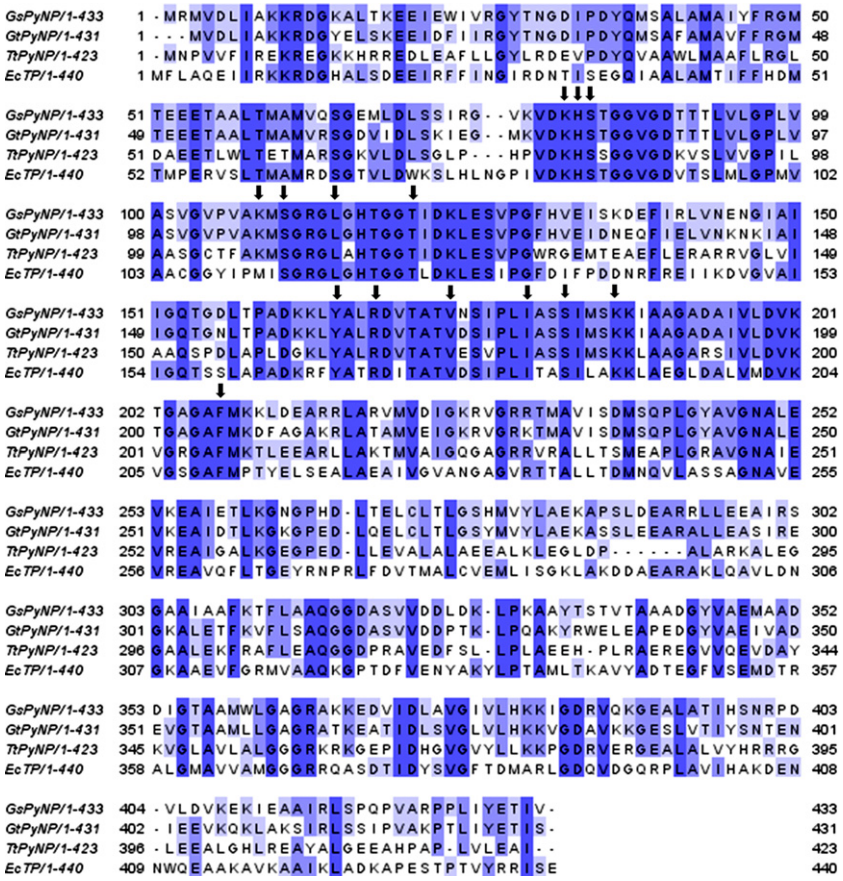


Fig. 1. 1Multiple sequence alignment of PyNP from *G. stearothermophilus* (GsPyNP), *G. thermoglucosidasius* (GtPyNP), *T. thermophilus* (TtPyNP) and TP from *E. coli* (EcTP). Shading represents the degree of sequence identity. Residues of the active site pocket are highlighted.

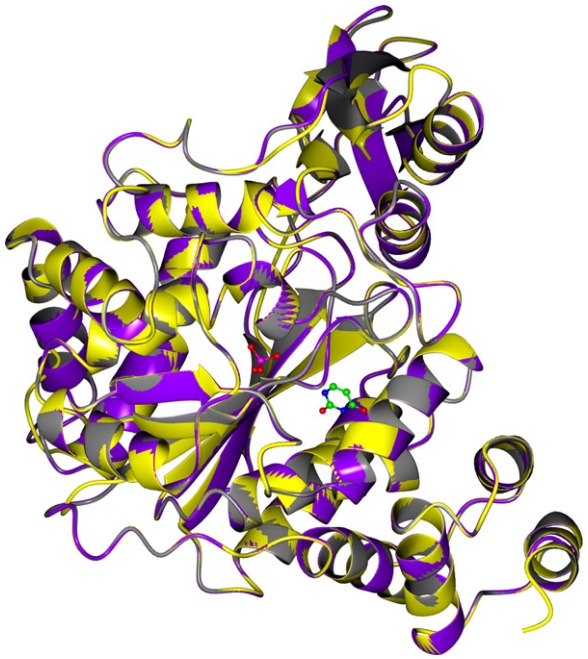


Fig. 2. 3D structure models. The superposition of GtPyNP (purple), TtPyNP (yellow) and GsPyNP (grey) 3D structures are shown. Structural models of GtPyNP and TtPyNP were built based on homology modeling using the structure of GsPyNP as template. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

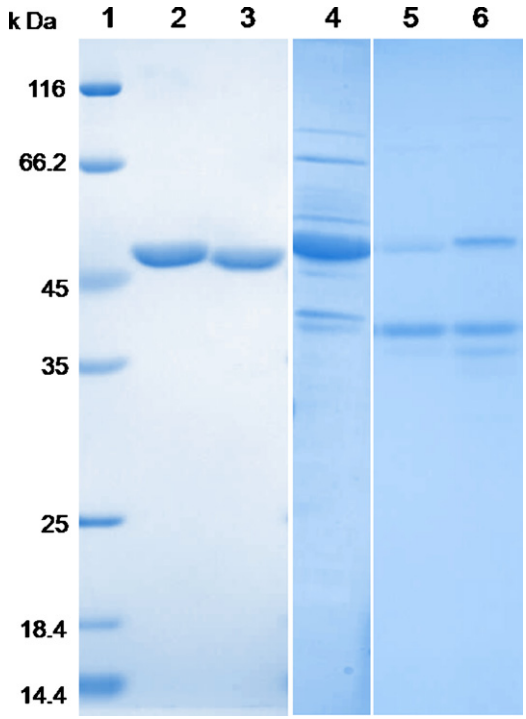


Fig. 3. SDS-PAGE of GtPyNP and TtPyNP samples. Recombinant expression of GtPyNP under standard conditions (37 °C, TB medium) was very efficient (4), but the expression of TtPyNP under the same conditions rather unsatisfactory (5). The expression level of TtPyNP could be significantly increased by using EnPresso medium that provides fed-batch like expression conditions (6). Purified GtPyNP (2) and TtPyNP (3) showed bands corresponding to the theoretically calculated molecular weight of the subunits (47.6 kDa and 46.9 kDa, respectively). Lane (1) represents the molecular weight marker.

acids that have been described to be involved in substrate binding or in the catalytic mechanism, respectively, are indicated [30,31].

The chain B of GsPyNP (pdb code 1BRW) was used as template to model both TtPyNP and GtPyNP. The structure of GsPyNP was elucidated in its closed conformation [30], uracil and the phosphate ion can be seen in the active site pocket. The structural fold of the models of GtPyNP and TtPyNP are almost identical to the template structure GsPyNP, which can be seen in the superposition of the secondary structure elements (Fig. 2).

3.2. Expression and purification

The desired sequences were expressed with an artificial N-terminal hexahistidine tag. In addition to a simple protein purification procedure this method offers the possibility to avoid expression problems associated with secondary structure formation of the 5′-end of mRNA, which is a critical issue, especially for genes derived from thermophiles [32]. Initially, the recombinant expression in *E. coli* of both GtPyNP and TtPyNP was performed under standard conditions (37 °C) in terrific broth (TB) medium. While the expression level of GtPyNP was satisfactory from the beginning (Fig. 3, lane 4), TtPyNP was expressed only very poorly (Fig. 3, lane 5). Hence, we tested also EnPresso medium for TtPyNP expression. This medium is based on an enzyme controlled substrate delivery and was reported to lead to higher cell densities and productivity per cell [33]. Indeed, the expression level of TtPyNP increased significantly (Fig. 3, lane 6) and the final cell density doubled (from $OD_{600} = 11$ to $OD_{600} = 22$). The volumetric yield obtained with this condition for TtPyNP expression was 0.06 mg ml^{-1}. Expression of GtPyNP in TB medium resulted in a volumetric yield of 0.17 mg ml^{-1}. Applied to SDS-PAGE gels, the purified proteins showed bands representing molecular weights that are consistent with the theoretically calculated values of the monomers (Fig. 3, lanes 2 and 3).

3.3. Thermal characteristics

The effect of temperature on the activity of the enzymes, as well as their thermal stability was investigated.

GtPyNP showed a temperature optimum of 60 °C, while the relative activity of TtPyNP increased with the reaction temperature, up to the highest temperature tested (99 °C, Fig. 4A). An apparent melting temperature of ≥102 °C and 103 °C was determined by circular dichroism and differential scanning calorimetry, respectively (Fig. 5). Hence, we assume that the temperature optimum of TtPyNP is in the range of 95–103 °C.

The stability half life of GtPyNP was estimated to be 1.6 h at 70 °C; while at 60 °C no significant loss of activity could be seen within 16 h of incubation. The stability half life of TtPyNP at 80 °C exceeds 23 h (Fig. 4B).

Compared to other reported enzymes with pyrimidine nucleoside phosphorylase activity, including UPs and TPs, the thermal stability and the temperature optimum of TtPyNP are extremely high (Table 1).

3.4. Kinetic parameters

Kinetic parameters describing the biocatalytical properties were determined at 60 °C for GtPyNP and at 80 °C for TtPyNP. The results are summarized in Table 2. The estimated K_m values of TtPyNP for the natural substrates uridine and thymidine are slightly lower than the K_m values determined by Hori et al. [1] for GsPyNP. In contrast to TtPyNP, GtPyNP is characterized by extremely low substrate affinities (high K_m values) towards both natural substrates.

The catalytic efficiency of an enzyme is best described by the ratio of k_{cat}/K_m [28]. These ratios are 15 times (substrate

Table 1
Thermal characteristics of reported PyNP, UP and TP.

Enzyme	Organism	Thermal stability ($t_{1/2}$)	Temp. optimum	Reference
TP	*E. coli*	<10 min (55 °C)	–	[34]
UP	*E. coli*	9.9 h (60 °C)	40 °C	[35]
PyNP	*G. thermoglucosidasius*	1.6 h (70 °C)	60 °C	This study
$UP_{engineered}$	*E. coli*	3.3 h (70 °C)	60 °C	[35]
UP	*Enterobacter aerogenes*	1 week (60 °C)	65 °C	[36]
PyNP	*G. stearothermophilus*	25 min (70 °C)	70 °C	[1,2]
UP	*Erwinia carotovora*	–	70 °C	[37]
PyNP	*T. thermophilus*	>24 h (80 °C)	>95 °C	This study

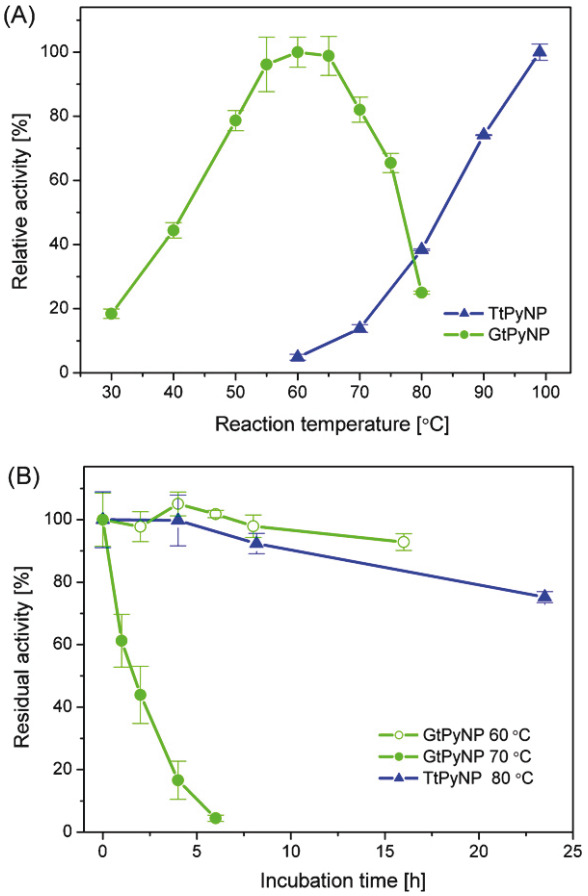


Fig. 4. Thermal characteristics. Relative activity of GtPyNP and TtPyNP over the reaction temperature (A), where the highest reaction rate determined was set to 100% for each enzyme. Thermostability (B) was investigated by incubating protein samples for defined time intervals and subsequently determining the residual activity, where the activity of enzyme samples that were not thermo-treated was set to 100%.

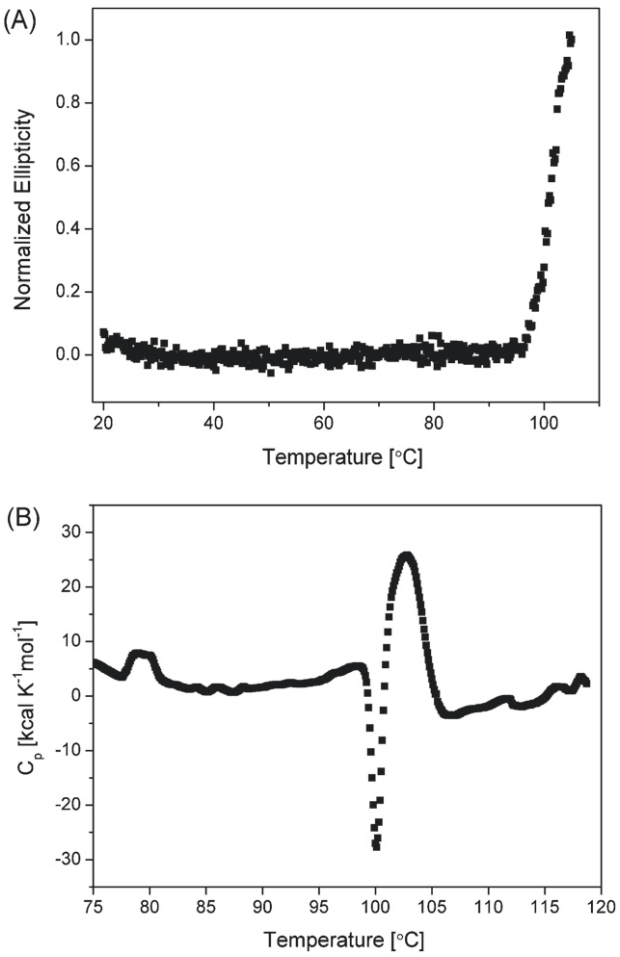


Fig. 5. Thermal unfolding of TtPyNP. Thermal unfolding trace monitored by loss of the far-UV circular dichroism signal at 220 nm provided an apparent melting temperature of at least 102 °C (A). The apparent melting temperature determined by differential scanning calorimetry (B) was 103 °C.

uridine) and 25 times (substrate thymidine) higher for TtPyNP than for GtPyNP. The k_{cat}/K_m ratio is also a measure to compare an enzyme's specificity towards different substrates [28]. The results of the present study show that both enzymes are more specific for uridine than for thymidine, but the difference in specificity is more pronounced for GtPyNP: the k_{cat}/K_m ratio is 2 fold higher for uridine than for thymidine. In contrast, the k_{cat}/K_m ratio of TtPyNP is only 1.23 fold higher for uridine vs that of thymidine.

Table 2
Kinetic parameters of reported PyNP in comparison to *E. coli* enzymes TP and UP.

Enzyme	K_m (μM)		k_{cat} (s^{-1})		k_{cat}/K_m (s^{-1} μM^{-1})		React. conditions	Reference
	Uridine	Thymidine	Uridine	Thymidine	Uridine	Thymidine		
GtPyNP	2342	1282	275	83	0.12	0.06	60 °C, pH 7.0	This study
TtPyNP	145	435	279	679	1.92	1.56	80 °C, pH 7.0	This study
GsPyNP	190	460	–	–	–	–	60 °C, pH 7.0	[1]
EcTP	60	300	$<1 \times 10^{-4}$	198	$<1.7 \times 10^{-6}$	0.66	25 °C, pH 6.5	[38,39]
EcUP	80	270	98	5	1.22	0.02	25 °C, pH 7.5	[38,40]

The estimation errors of the K_m and k_{cat} values determined in this study were not higher than 10%. (–) Data not indicated.

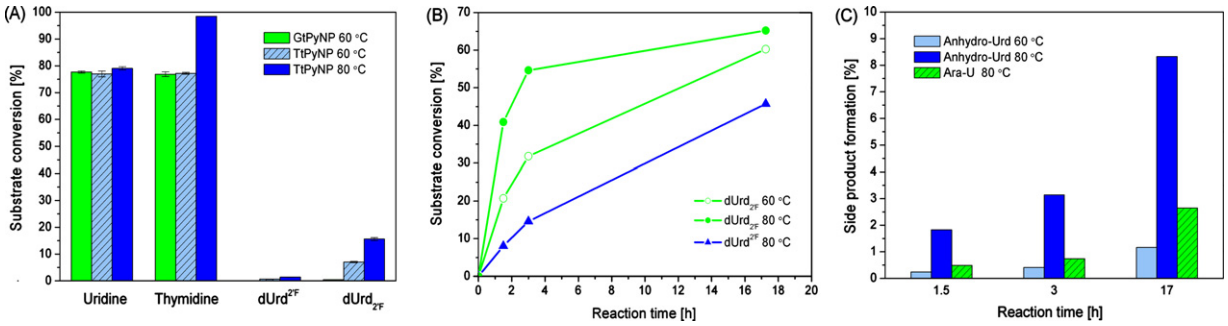


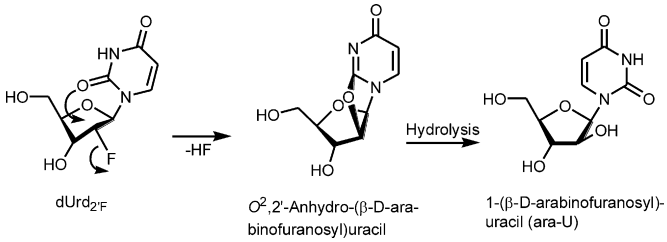
Fig. 6. Phosphorolysis of natural and modified pyrimidine nucleosides. (A) The percentage of natural and 2′-fluorosubstituted pyrimidine nucleosides that were phosphorolytically cleaved by TtPyNP or GtPyNP after 30 min, (B) percentage of 2′-fluorosubstituted pyrimidine nucleosides phosphorolytically cleaved by TtPyNP after prolonged reaction times at indicated temperatures and (C) percentage of $dUrd_{2'F}$ molecules that reacted to anhydro-Urd and ara-U during the TtPyNP catalyzed phosphorolytic cleavage reaction in dependence of reaction time and temperature. The ara-U formation at 60 °C could not be accurately determined but was estimated not to be higher than 0.4%.

In synthetic applications, where high substrate concentrations are used ($c_s \gg K_m$) the k_{cat} value may be the most appropriate parameter describing the efficiency of the biocatalyst. The turnover numbers (k_{cat}) of GtPyNP and TtPyNP are in similar range, with uridine as substrate. By contrast, the turnover numbers for thymidine differ significantly, in favor of thymidine phosphorolysis by TtPyNP. The k_{cat} value of TtPyNP for thymidine is also unusually high in comparison to alternative enzymes that are used for the phosphorolysis of pyrimidine nucleosides, e.g. EcUP and EcTP (Table 2).

3.5. Phosphorolysis of 2′-fluoro substituted pyrimidine nucleosides

Of particular interest is the potential of both thermostable PyNPs as biocatalysts in the synthesis of modified nucleosides. With this aim in view, the phosphorolysis of natural pyrimidine nucleoside substrates (thymidine and uridine) and their sugar modified analogues, *viz.*, 2′-deoxy-2′-fluorouridine ($dUrd_{2'F}$) and 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)uracil ($dUrd^{2'F}$), were investigated. These substrates can be used as pentofuranosyl donors in enzymatic transglycosylations aiming at the synthesis of pharmaceutically valuable 2′-fluorosubstituted purine nucleosides. With this strategy $dUrd_{2'F}$ served as a substrate for the enzymatic synthesis of 2′-deoxy-2′-fluoroguanosine using the whole *E. coli* cells as a biocatalyst [41] and a multitude of other purine 2′-deoxy-2′-fluororibosides with antiviral activity using a combination of EcTP and EcPNP as a biocatalyst [42].

However, $dUrd_{2'F}$ and $dUrd^{2'F}$ are very poor substrates in phosphorolysis reactions. This is presumably a result of increased strength of the glycosyl bond as it follows from the crystallographic data for the N1–C1′ bond length of uridine (aver. value 1.490 Å [43]) and its 2′-deoxyfluoro analogues (1.454 Å [44] and 1.460 Å [45], respectively). Moreover, introduction of a fluorine atom into pentofuranose ring of nucleosides results in dramatic changes of the conformation of such analogues precluding the formation of the productive substrate-catalytic center complex (for more detailed discussion, see [11]).



Scheme 3.

Indeed, it was reported that (i) $dUrd_{2'F}$ showed no detectable substrate activity towards EcUP, (ii) EcTP catalyzed the phosphorolysis of $dUrd_{2'F}$ but at an extremely low rate, and (iii) the enzymatic cleavage of the glycosidic bond of $dUrd^{2'F}$ equally afforded a high amount of enzyme and prolonged reaction time (6 days) [42,46].

In this study we have investigated the phosphorolysis of these challenging substrates by GtPyNP and TtPyNP (Fig. 6). Our results indicate that TtPyNP might be a good alternative to the use of *E. coli* enzymes, but the use of GtPyNP is apparently not suitable for the applications discussed above: no activity towards $dUrd^{2'F}$ and only poor activity towards $dUrd_{2'F}$ (0.44% substrate conversion) was detected with GtPyNP as biocatalyst after 30 min reaction time.

By contrast, the TtPyNP catalyzed reaction under the same conditions resulted in the phosphorolytic cleavage of 0.65% of $dUrd^{2'F}$ and 7.0% of $dUrd_{2'F}$. Since the optimal reaction temperature of TtPyNP is significantly higher than 60 °C (see Section 3.3), we repeated the same reaction also at 80 °C. Now, the TtPyNP catalyzed reaction resulted in 1.4% phosphorolyzed $dUrd^{2'F}$ and 15.6% of $dUrd_{2'F}$. However, under these conditions the formation of two new peaks were observed by HPLC analysis of the reaction mixture that contained $dUrd_{2'F}$. The retention times coincide with those of authentic samples of O^2,2′-anhydro-1-(β-D-arabinofuranosyl)uracil (anhydro-Urd) (2.3 min) and 1-(β-D-arabinofuranosyl)uracil (ara-U) (3.8 min). Hence, the formation of anhydro-Urd resulting from HF release from the $dUrd_{2'F}$ molecule and the subsequent hydrolysis of anhydro-Urd resulting in ara-U appear to be a reasonable explanation (Scheme 3). Phosphorolysis of both 2′-fluorosubstituted pyrimidine nucleosides catalyzed by TtPyNP was also monitored over prolonged reaction times (Fig. 6B). The results show that the conversion of $dUrd^{2'F}$ at 80 °C could be increased to 46% after 17 h; side-product formation was not observed. The final amount of phosphorolyzed $dUrd_{2'F}$ after 17 h at 80 °C (65%) was in similar range as the amount obtained at 60° (60%). By contrast, side product formation, as discussed above, at 80 °C was significantly higher than at 60 °C (Fig. 6C). After 17 h 8.3% of $dUrd_{2'F}$ reacted to anhydro-Urd at 80 °C, whereas only 1.2% of $dUrd_{2'F}$ reacted to anhydro-Urd at 60 °C.

4. Conclusion

PyNPs isolated from thermophilic microorganisms are promising biocatalysts for the efficient synthesis of modified nucleosides [18]. Up to now, the biocatalytic characterization of PyNPs from thermophilic microbes was restricted to PyNPs from *G. stearothermophilus* strains [1,2,6]. We have studied here the biocatalytical properties of two additional thermostable PyNP, originating from *G. thermoglucosidasius* and *T. thermophilus.*

Our results indicate that both enzymes show excellent biocatalytical properties for applications with natural pyrimidine nucleosides as substrates (thymidine, uridine) and a reaction temperature of 60 °C. In addition, the unusually high thermal stability of TtPyNP makes this biocatalyst also suitable for reactions requiring an even higher reaction temperature of 80 °C.

We have further tested both thermostable PyNPs towards their ability to phosphorolyze 2′-fluorosubstituted pyrimidine nucleosides that have been shown to be very poor substrates in phosphorolysis reactions employing EcUP or EcTP as biocatalyst [42,46]. Our results reveal striking differences of the substrate specificities of GtPyNP and TtPyNP, in favor of the latter. These findings make TtPyNP a candidate as powerful biocatalyst in the transglycosylation reactions aiming at the synthesis of 2′-fluoro substituted purine nucleosides. Expression of appropriate PNPs and their use in synthetic applications are ongoing.

Acknowledgments

This work is part of the Cluster of Excellence "Unifying Concepts in Catalysis" coordinated by the Technische Universität Berlin. Financial support by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the German Initiative for Excellence is gratefully acknowledged (EXC 314). We would like to thank Prof. Reinhard Sterner (Institute of Biophysics and Physical Biochemistry, University of Regensburg) for his support and providing us the facilities for protein analytical measurements. 1-(2-Deoxy-2-fluoro-β-D-arabinofuranosyl)uracil was kindly provided by Prof. Alex Azhayev (Metkinen Chemistry). We further thank Prof. Xiaoda Yang (School of Pharmaceutical Sciences, Peking University Health Science Centre), Dr. Andy Maraite and Prof. Marion Ansorge-Schumacher (Technische Universität Berlin) for fruitful discussions.

References

[1] N. Hori, M. Watanabe, Y. Yamazaki, Y. Mikami, Agric. Biol. Chem. 54 (1990) 763–768.
[2] T. Hamamoto, T. Noguchi, Y. Midorikawa, Biosci. Biotechnol. Biochem. 60 (1996) 1179–1180.
[3] M.J. Pugmire, S.E. Ealick, J. Biochem. 361 (2002) 1–25.
[4] T. Furukawa, A. Yoshimura, T. Sumizawa, M. Haraguchi, S. Akiyama, K. Fukui, M. Ishizawa, Y. Yamada, Nature 356 (1992) 668.
[5] S. Akiyama, T. Furukawa, T. Sumizawa, Y. Takebayashi, Y. Nakajima, S. Shimaoka, M. Haraguchi, Cancer Sci. 95 (2004) 851–857.
[6] P.P. Saunders, B.A. Wilson, G.F. Saunders, J. Biol. Chem. 244 (1969) 3691–3697.
[7] M.G. Vander Heiden, Nat. Rev. Drug Discov. 10 (2011) 671–684.
[8] E. De Clercq, Annu. Rev. Pharmacol. 51 (2011) 1–24.
[9] A.J. Berdis, Biochemistry 47 (2008) 8253–8260.
[10] P. Herdewijn, Modified Nucleosides in Biochemistry, Biotechnology and Medicine, Wiley-VCH, Weinheim, 2008.
[11] I.A. Mikhailopulo, A.I. Miroshnikov, Mendeleev Commun. 21 (2011) 57–68.
[12] N. Hori, M. Watanabe, Y. Yamazaki, Y. Mikami, Agric. Biol. Chem. 53 (1989) 2205–2210.
[13] A. Bzowska, E. Kulikowska, D. Shugar, Pharmacol. Ther. 88 (2000) 349–425.
[14] G. Cacciapuoti, S. Gorassini, M.F. Mazzeo, R.A. Siciliano, V. Carbone, V. Zappia, M. Porcelli, FEBS J. 274 (2007) 2482–2495.
[15] D. Visser, F. Hennessy, K. Rashamuse, M. Louw, D. Brady, Extremophiles 14 (2010) 185–192.
[16] X.F. Gao, X.R. Huang, C.C. Sun, J. Struct. Biol. 154 (2006) 20–26.
[17] K. Shimizu, N. Kunishima, Acta Crystallogr. F 63 (2007) 308–310.
[18] S.A. Taran, K.N. Verevkina, S.A. Feofanov, A.I. Miroshnikov, Russ. J. Bioorg. Chem. 35 (2009) 739–745.
[19] T. Schwab, R. Sterner, Chembiochem 12 (2011) 1581–1588.
[20] D. Moore, D. Dowhan, in: F.M. Ausubel, et al. (Eds.), Purification and Concentration of DNA from Aqueous Solutions, John Wiley & Sons, Houston, 2002, Unit 2.1A.
[21] J. Šiurkus, J. Panula-Perälä, U. Horn, M. Kraft, R. Rimšeliene, P. Neubauer, Microb. Cell Fact. 9 (2010) 35.
[22] K. Gerdes, Biotechnology 6 (1988) 1402–1405.
[23] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. 215 (1990) 403–410.
[24] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Bioinformatics 23 (2007) 2947–2948.
[25] L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey, T. Schwede, Nat. Protoc. 4 (2009) 1–13.
[26] S. McNicholas, E. Potterton, K.S. Wilson, M.E.M. Noble, Acta Crystallogr. D 67 (2011) 386–394.
[27] J. Sambrook, D.W. Russell, Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
[28] R.A. Copeland, ENZYMES a Practical Introduction to Structure, Mechanism, and Data Analysis, 2nd ed, Wiley-VCH, New York, 2000.
[29] A.A. Akhrem, G.V. Zaitseva, I.A. Mikhailopulo, A.F. Abramov, J. Carbohydr. Nucl. 4 (1977) 43–63.
[30] M.J. Pugmire, S.E. Ealick, Structure 6 (1998) 1467–1479.
[31] J. Mendieta, S. Martin-Santamaria, E.M. Priego, J. Balzarini, M.J. Camarasa, M.J. Perez-Perez, F. Gago, Biochemistry 43 (2004) 405–414.
[32] K. Szeker, O. Niemitalo, M.G. Casteleijn, A.H. Juffer, P. Neubauer, J. Biotechnol. 156 (2010) 268–274.
[33] K. Ukkonen, A. Vasala, H. Ojamo, P. Neubauer, Microb. Cell Fact. 10 (2011) 107.
[34] T.A. Krenitsky, J.V. Tuttle, Biochim. Biophys. Acta 703 (1982) 247–249.
[35] D.F. Visser, F. Hennessy, J. Rashamuse, B. Pletschke, D. Brady, J. Mol. Catal. B: Enzym. 68 (2011) 279–285.
[36] T. Utagawa, H. Morisawa, Y. Shigeru, A. Yamazaki, F. Yoshinaga, Y. Hirose, Agric. Biol. Chem. 49 (1985) 3239–3246.
[37] H. Shirae, K. Yokozeki, Agric. Biol. Chem. 55 (1991) 1849–1857.
[38] N.G. Panova, E.V. Shcheveleva, K.S. Alekseev, V.G. Mukhortov, A.N. Zuev, S.N. Mikhailov, R.S. Esipov, D.V. Chuvikovskii, A.I. Miroshnikov, Mol. Biol. (Mosk) 38 (2004) 907–913.
[39] N.G. Panova, C.S. Alexeev, A.S. Kuzmichov, E.V. Shcheveleva, S.A. Gavryushov, K.M. Polyakov, A.M. Kritzyn, S.N. Mikhailov, R.S. Esipov, A.I. Miroshnikov, Biochemistry (Moscow) 72 (2007) 21–28.
[40] C.S. Alexeev, N.G. Panova, K.M. Polyakov, S.N. Mikhailov, Abstracts of XIX International Round Table on Nucleosides, Nucleotides and Nucleic Acids IRT 2010, 2010, p. 94.
[41] G.V. Zaitseva, A.I. Zinchenko, V.N. Barai, N.I. Pavlova, E.I. Boreko, I.A. Mikhailopulo, Nucleosides Nucleotides 18 (1999) 687–688.
[42] J.V. Tuttle, M. Tisdale, T.A. Krenitsky, J. Med. Chem. 36 (1993) 119–125.
[43] E.A. Green, R.D. Rosenstein, R. Shiono, D.J. Abraham, B.L. Trus, R.E. Marsh, Acta Crystallogr. B 31 (1975) 102–107.
[44] C. Marck, B. Lesyng, W. Saenger, J. Mol. Struct. 82 (1982) 77–94.
[45] A. Hempel, N. Camerman, J. Grierson, D. Mastropaolo, A. Camerman, Acta Crystallogr. C 55 (1999) 632–633.
[46] J.V. Tuttle, T.A. Krenitsky, European Patent Publication EP 0285432 B1 (1992).

Paper II

the FEBS Journal

Recombinant purine nucleoside phosphorylases from thermophiles: preparation, properties and activity towards purine and pyrimidine nucleosides

Xinrui Zhou[1], Kathleen Szeker[1], Bernd Janocha[1,2], Thomas Böhme[1,2], Dirk Albrecht[3], Igor A. Mikhailopulo[4] and Peter Neubauer[1]

1 Laboratory of Bioprocess Engineering, Department of Biotechnology, Technische Universität Berlin, Germany
2 Sanofi-Aventis Deutschland GmbH, Frankfurt, Germany
3 Institute of Microbiology, Ernst-Moritz-Arndt-Universität Greifswald, Germany
4 Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Minsk, Belarus

Keywords
2′–deoxyfluoro adenine nucleosides; *Aeropyrum pernix*; *Deinococcus geothermalis*; *Geobacillus thermoglucosidasius*; thermostable PNP

Correspondence
P. Neubauer, Laboratory of Bioprocess Engineering, Department of Biotechnology, Technische Universität Berlin, Ackerstr. 71–76, ACK24, D–13355 Berlin, Germany
Fax: +49 3031 427577
Tel: +49 3031 472269
E-mail: peter.neubauer@tu-berlin.de
Website: http://www.bioprocess.tu-berlin.de
I. A. Mikhailopulo, Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Acad. Kuprevicha 5/2, 220141 Minsk, Belarus
Fax: +375 1726 78161
Tel: +375 1726 78148
E-mail: imikhailopulo@gmail.com

(Received 6 December 2012, revised 13 January 2013, accepted 16 January 2013)

doi:10.1111/febs.12143

Thermostable nucleoside phosphorylases are attractive biocatalysts for the synthesis of modified nucleosides. Hence we report on the recombinant expression of three 'high molecular mass' purine nucleoside phosphorylases (PNPs) derived from the thermophilic bacteria *Deinococcus geothermalis*, *Geobacillus thermoglucosidasius* and from the hyperthermophilic archaeon *Aeropyrum pernix* (5′–methythioadenosine phosphorylase; ApMTAP). Thermostability studies, kinetic analysis and substrate specificities are reported. The PNPs were stable at their optimal temperatures (DgPNP, 55 °C; GtPNP, 70 °C; ApMTAP, activity rising to 99 °C). Substrate properties were investigated for natural purine nucleosides [adenosine, inosine and their C2′-deoxy counterparts (activity within 50–500 $U \cdot mg^{-1}$)], analogues with 2′-amino modified 2′-deoxy-adenosine and -inosine (within 0.1–3 $U \cdot mg^{-1}$) as well as 2′-deoxy-2′-fluoroadenosine (**9**) and its C2′-*arabino* diastereomer (**10**, within 0.01–0.03 $U \cdot mg^{-1}$). Our results reveal that the structure of the heterocyclic base (e.g. adenine or hypoxanthine) can play a critical role in the phosphorolysis reaction. The implications of this finding may be helpful for reaction mechanism studies or optimization of reaction conditions. Unexpectedly, the diastereomeric 2′-deoxyfluoro adenine *ribo*- and *arabino*-nucleosides displayed similar substrate properties. Moreover, cytidine and 2′-deoxycytidine were found to be moderate substrates of the prepared PNPs, with substrate activities in a range similar to those determined for 2′-deoxyfluoro adenine nucleosides **9** and **10**. C2′-modified nucleosides are accepted as substrates by all recombinant enzymes studied, making these enzymes promising biocatalysts for the synthesis of modified nucleosides. Indeed, the prepared PNPs performed well in preliminary transglycosylation reactions resulting in the synthesis of 2′-deoxyfluoro adenine *ribo*- and *arabino*- nucleosides in moderate yield (24%).

Abbreviations
Ade, adenine; Ado, adenosine; Ap, *Aeropyrum pernix*; Cyd, cytidine; dAdo, 2′–deoxyadenosine; $dAdo_{2'F}$, 2′–deoxy-2′–fluoroadenosine; $dAdo^{2'F}$, 9–(2-deoxy-2-fluoro- -D-arabinofuranosyl)adenine; $dAdo_{2'NH2}$, 2′–amino-2′-deoxyadenosine; dCyd, 2′–deoxycytidine; Dg, *Deinococcus geothermalis*; dIno, 2′–deoxyinosine; $dIno_{2'NH2}$, 2′–amino-2′–deoxyinoinosine; $dUrd^{2'F}$, 1–(2-deoxy-2-fluoro- -D-arabinofuranosyl) uracil; $dUrd_{2'F}$, 2′–deoxy-2′–fluoroadenosine; Ea, *Enterobacter aerogenes*; Ec, *Escherichia coli*; Gs, *Geobacillus stearothermophilus*; Gt, *Geobacillus thermoglucosidasius*; Ino, inosine; IPTG, isopropyl β-D–thiogalactoside; MTAP, 5′–methythioadenosine phosphorylase; NP, nucleoside phosphorylase; Pf, *Pyrococcus furiosus*; PNP, purine nucleoside phosphorylase; PyNP, pyrimidine nucleoside phosphorylase; Ss, *Sulfolobus solfataricus*; Tt, *Thermus thermophilus*; Urd, uridine.

Introduction

Base and sugar-modified analogues of natural nucleosides are widely used as pharmaceutical agents for the treatment of viral infections and cancer [1–3]. Moreover, the replacement of natural nucleosides by modified species results in improvement in the biochemical properties of synthetic oligonucleotides. For example, C2′–fluorinated nucleosides have been shown to have favorable properties as components of antisense oligonucleotides and small interfering RNAs [4–6]. The chemical synthesis of modified nucleosides is challenged by stereo- and regioselective requirements, as well as the need to protect and deprotect sensitive functional groups. By contrast, enzymes, such as nucleoside phosphorylases (NPs) can be used as biocatalysts for the efficient synthesis of nucleoside analogues under environmentally friendly conditions [7–11].

However, common disadvantages encountered in biocatalysis comprise the instability of many enzymes under harsh reaction conditions and narrow substrate spectra limiting the scope of any potential synthetic reactions. *Escherichia coli* NP-mediated synthesis of fluorinated nucleosides represents an example of this dilemma [12]. The use of enzymes from thermophiles is a reasonable strategy for overcoming instability problems [13]. Such thermozymes offer the possibility of performing reactions at high temperature, resulting in diminution of the viscosity of the medium and increased substrate diffusion coefficients, and leads to higher overall reaction rates. The final yield may be improved because of the higher solubility of substrates with the increase in temperature [14]. It is noteworthy that thermozymes are often resistant to pressure and organic solvents [15,16]. They can be expressed in *E. coli* at high levels and are easily purified by heat treatment [17,18].

Purine nucleoside phosphorylase (PNP; EC 2.4.2.1) and 5′–methylthioadenosine phosphorylase (MTAP; EC 2.4.2.28) catalyze the reversible phosphorolysis of purine nucleosides, as depicted in Scheme 1. Both belong to the PNP family [19]. According to Shugar and co-workers, PNPs can be grouped into two main classes (for a comprehensive review, see ref. [19]): PNPs of the first class (‘low molecular mass PNPs’), found mainly in mammalian tissues and in some bacteria, are specific for 6–oxo purine nucleosides only; those of the second class (‘high molecular mass PNPs’) are predominantly found in bacteria and accept 6–oxo (e.g. inosine and guanosine) as well as 6–amino (e.g. adenosine) purine nucleosides as substrates. Because of the broader substrate specificity, the latter group of enzymes is especially interesting for synthetic applications. Although many PNPs have been extensively studied [19], examples of ‘high molecular mass PNPs’ derived from thermophiles are limited to those from *Geobacillus stearothermophilus* (GsPNP, *punB* gene) [9,20] and *Enterobacter aerogenes* (EaPNP) [21,22]; the MTAP from *Aeropyrum pernix* (ApMTAP) [23], *Sulfolobus solfataricus* (SsMTAP) [24,25] and *Pyrococcus furiosus* (PfMTAP) [26], whereby only GsPNP, EaPNP and ApMTAP were considered for synthetic applications. Besides, ApMTAP has been recently applied in synthesis reactions [23], however, the information about the enzyme’s characteristics is not available.

In this study, we report on: (a) the preparation of three recombinant ‘high molecular mass’ PNPs derived from the thermophilic bacteria *Deinococcus geothermalis* (DgPNP) and *Geobacillus thermoglucosidasius* 11955 (GtPNP), as well as from the hyperthermophilic archaeon *A. pernix* (ApMTAP); and (b) the investigation into their physicochemical and substrate specificity. Here, we report on their thermal properties, kinetic parameters and substrate specificities in regard to natural purine substrates (inosine, adenosine and their 2′–deoxy counterparts, **1** and **2**, Scheme 1), analogues of natural substrates modified at the C2′ carbon atom [2′–amino-2′–deoxyadenosine (**7**), 2′–amino-2′–deoxyinosine (**8**), 2′–deoxy-2′–fluoroadenosine (**9**) and its C2′–*arabino* diastereomer **10** (Scheme 2)].

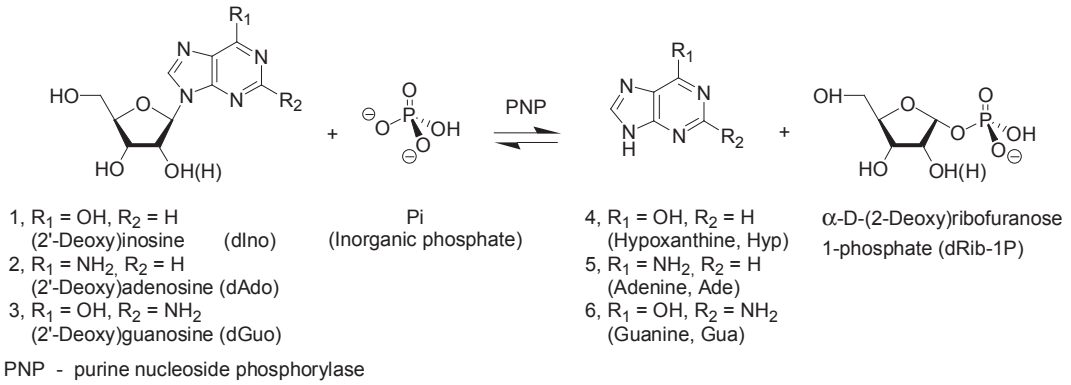
Our results reveal that 2′–modified nucleosides are accepted as substrates by all recombinant enzymes studied, indicating that these enzymes are promising biocatalysts for the synthesis of their respective nucleoside analogues. Indeed, in preliminary experiments for the synthesis of 2′–deoxyfluoro adenine nucleosides **9** and **10**, when the prepared PNPs were applied in combination with thermostable pyrimidine nucleosides phosphorylases (PyNP) [17], using adenine as an acceptor and the respective 2′–deoxy-2′–fluorouridine ($dUrd_{2'F}$) and 1-(2-deoxy-2-fluoro-β-D–arabinofuranosyl)uracil ($dUrd^{2'F}$) as the pentofuranose residue donors, the formation of the desired nucleosides was observed with a yield of ~ 24%.

Finally, we tested cytidine (**11**) and 2′–deoxycytidine (**12**) as substrates of the PNPs prepared and found that these nucleosides showed moderate substrate activity compared with that of 2′–deoxyfluoro adenine nucleosides **9** and **10**.

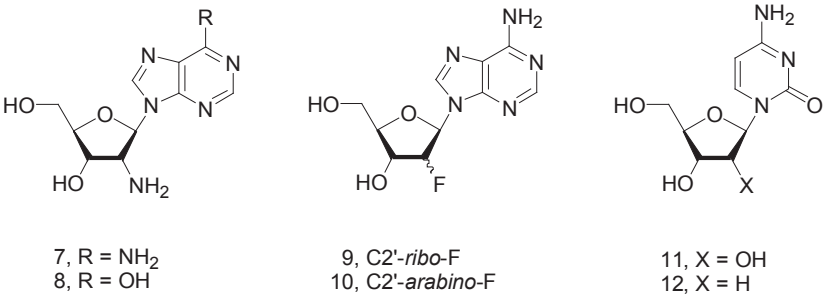
Results

Cloning, expression and purification

In order to avoid problems arising from stable secondary structures in the 5′–mRNA region [27], the target



Scheme 1. Phosphorolysis of purine nucleosides by PNP.



Scheme 2. Non-natural substrates of PNPs in this study.

gene sequences were cloned with an N–terminal hexahistidine tag.

All three enzymes were functionally expressed in *E. coli* BL21. Soluble GtPNP was well expressed under standard conditions [TB medium, 37 °C, 100 μM isopropyl β–D–thiogalactoside (IPTG)] as shown in Fig. 1A (lane 1). For the expression of DgPNP (Fig. 1A, lane 3) and ApMTAP (Fig. 1A, lane 5), EnPresso® medium [28] was used at a lower IPTG concentration (20 μM) to prevent the formation of insoluble protein aggregates. The final cell densities varied from A_{600} = 19 (ApMTAP) to 25 (DgPNP) and 42 (GtPNP).

In SDS/PAGE analysis, the purified proteins (Fig. 1A, lanes 2, 4 and 6) showed bands corresponding to the theoretically calculated molecular masses of the monomeric subunits (DgPNP, 30.0 kDa; GtPNP, 27.6 kDa; and ApMTAP, 27.0 kDa).

However, in the samples of purified ApMTAP, a second band with a molecular mass of ~ 116 kDa could be seen. It was confirmed by MS analysis that this second band also represents ApMTAP (Tables S1 and S2). Because disulfide bond formation is considered to be a key factor for protein stabilization in thermophilic archaea [29–31], we assumed that the second band related to one of possible oligomerization states of the ApMTAP monomeric subunits connected by the disulfide bonds. This assumption was supported by SDS/PAGE analysis (Fig. 1B, lanes 3 and 6), which revealed that the oligomeric band was predominant when the purified ApMTAP was heat-treated in the absence of dithiothreitol (Fig. 1B, lane 3), whereas the monomeric band was predominant after dithiothreitol treatment (Fig. 1B, lane 6). This phenomenon was not observed for the other proteins. Furthermore, we found that the higher oligomeric state of ApMTAP was catalytically active, whereas the monomeric ApMTAP was inactive (data not shown).

Thermal characteristics

The effect of temperature on the activity of the enzymes was examined between 30 and 99 °C (Fig. 2). DgPNP showed optimal activity at 55 °C (Fig. 2), which is slightly higher than the optimal growth temperature (45–50 °C) of *D. geothermalis* [32]. The stability half-life at 60 °C was 1.6 h, whereas 80% of the activity was retained at 55 °C for 8 h (Fig. 3A). GtPNP showed an optimal activity at 70 °C (Fig. 2), which is ~ 15 °C higher than the optimal growth

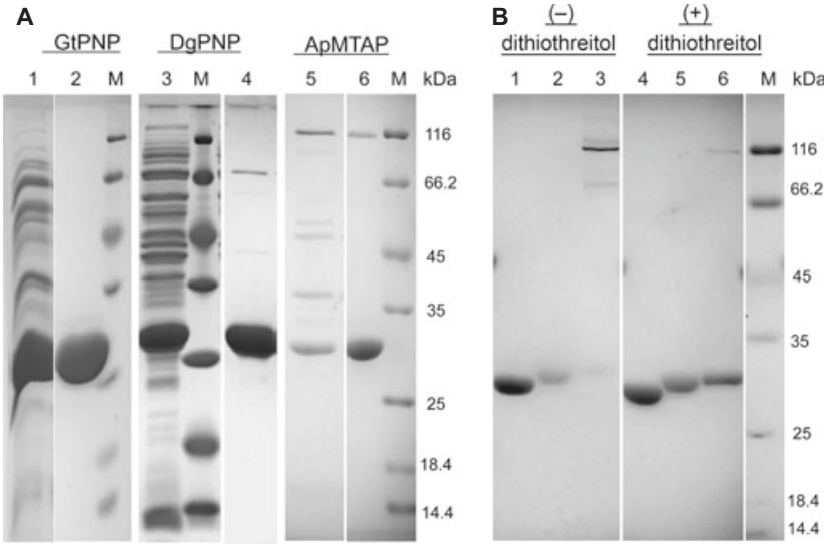


Fig. 1. SDS/PAGE analysis of recombinant PNPs overexpressed in *E. coli* BL21 (A) and the influence of dithiothreitol treatment on the enzymes (B). (A) Soluble fraction after incubation of the cell lysis at 65 °C (lane 1), 55 °C (lane 3) and 85 °C (lane 5) for 15 min. Purified samples of GtPNP (lane 2), DgPNP (lane 4) and ApMTAP (lane 6) after Ni ion-affinity chromatography and a desalting step. (B) SDS/PAGE samples of purified enzymes without dithiothreitol treatment GtPNP (lane 1), DgPNP (lane 2) and ApMTAP (lane 3); dithiothreitol treatment for purified GtPNP (lane 4), DgPNP (lane 5) and ApMTAP (lane 6). M, molecular mass marker.

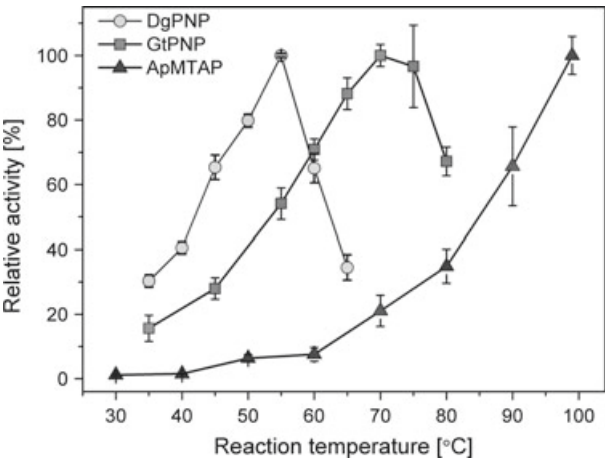


Fig. 2. Temperature dependence of the activity of DgPNP, GtPNP and ApMTAP. The highest reaction rate observed for each enzyme was set to 100%. Substrate: 1 mM inosine in 50 mM phosphate buffer (pH 7).

temperature (55 °C) of its source microorganism *G. thermoglucosidasius* [33] and 10 °C higher than the temperature optimum of GtPyNP, another NP enzyme, from the same microorganism [17]. GtPNP was incubated at 70 °C for 8 h and no significant activity loss could be observed (Fig. 3B), whereas GtPyNP at 70 °C had a half-life of only 1.6 h [17]. The activity of ApMTAP increased exponentially at increased reaction temperatures up to the highest temperature tested (99 °C, Fig. 2), while the source microorganism *A. pernix* shows optimal growth at temperatures between 90 and 95 °C [34]. ApMTAP is extremely thermostable with an estimated half-life of 69 h at 90 °C (Fig. 3C).

The thermal characteristics of the PNPs studied here along with other thermostable PNPs characterized by a broad substrate specificity (recognition of both 6–amino and 6–oxo purine nucleosides) are summarized in Table 1.

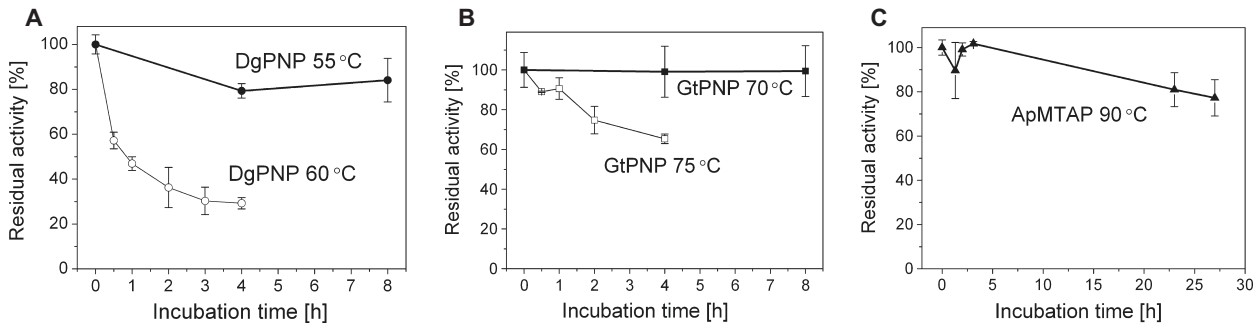


Fig. 3. Thermal stabilities of DgPNP (A), GtPNP (B) and ApMTAP (C). Enzymes were incubated at different temperatures and the residual activity after defined time intervals was determined under standard assay conditions (substrate: 1 mM inosine in 50 mM phosphate buffer, pH 7). Enzyme activities before incubation were set to 100%.

Table 1. Thermal characteristics of PNPs.

Enzyme	Organism	Thermal stability ($t_{1/2}$)	Temp. optimum [°C]	Ref.
EcPNP	*Escherichia coli*	> 30 min (50 °C); < 30 min (55 °C)	60	[35]
DgPNP	*Deinococcus geothermalis*	> 8 h (55 °C); 1.7 h (60 °C)	55	This study
EaPNP	*Enterobacter aerogenes*	7 days (60 °C)[a]	60	[21]
GtPNP	*Geobacillus thermoglucosidasius*	> 8 h (70 °C); 6.3 h (75 °C)	70	This study
GsPNP	*Geobacillus stearothermophilus*	> 24 h (70 °C)	70	[9,20]
ApMTAP	*Aeropyrum pernix*	69.3 h (90 °C)	> 99	This study
SsMTAP	*Sulfolobus solfataricus*	> 2 h (100 °C); 15 min (130 °C)	120	[25]
PfMTAP	*Pyrococcus furiosus*	> 5 h (100 °C); 43 min (130 °C)	125	[26]

[a] Purified enzyme was kept in 30 mM inosine (in 25 mM phosphate buffer) at 60 °C.

Substrate specificity

The substrate specificities of the enzymes prepared were studied using a number of natural purine nucleosides [adenosine (Ado), inosine (Ino), 2′–deoxyadenosine (dAdo), 2′–deoxyinosine (dIno)] and 2′-modified analogues ($dAdo_{2'NH2}$ and $dIno_{2'NH2}$, $dAdo_{2'F}$ and $dAdo^{2'F}$).

Phosphorolysis of purine nucleosides

The results show that all three enzymes recognize both 6–oxopurine (e.g. Ino, dIno, $dIno_{2'NH2}$) and 6–aminopurine nucleosides (e.g. Ado, dAdo, $dAdo_{2'NH2}$) (Fig. 4). Although specific activities for the same substrate are distinct for each enzyme, it can be generally stated that natural substrates are significantly better accepted than 2′–amino modified substrates (50–500 versus 0.05–3 $U \cdot mg^{-1}$), whereas the latter are significantly better accepted than the 2′–fluoro substituted compounds (0.01–0.03 $U \cdot mg^{-1}$). Moreover, it is noteworthy that ApMTAP is characterized by relatively low specific activities with natural substrates, but by a relatively broad substrate spectrum at the same time, including $dAdo_{2'NH2}$ and $dIno_{2'NH2}$, $dAdo_{2'F}$ and $dAdo^{2'F}$; whereas GtPNP is highly specific for natural substrates but also accepts modified substrates. DgPNP was the best candidate for phosphorolysis of the 2′–fluorinated purine nucleosides (Fig. 4). Interestingly, GtPNP and ApMTAP displayed similar levels of activity with $dAdo_{2'NH2}$. However, the replacement of the heterocyclic base adenine of $dAdo_{2'NH2}$ with hypoxanthine (resulting in $dIno_{2'NH2}$) led to a significant decrease in substrate activity in the case of GtPNP, whereas the substrate activity towards ApMTAP was slightly enhanced relative to $dAdo_{2'NH2}$ (Fig. 4). Notably, a similar characteristic of the GtPNP activity is observed in the case of reactions with Ado and Ino, whereas dAdo and dIno showed reversed dependence of activities.

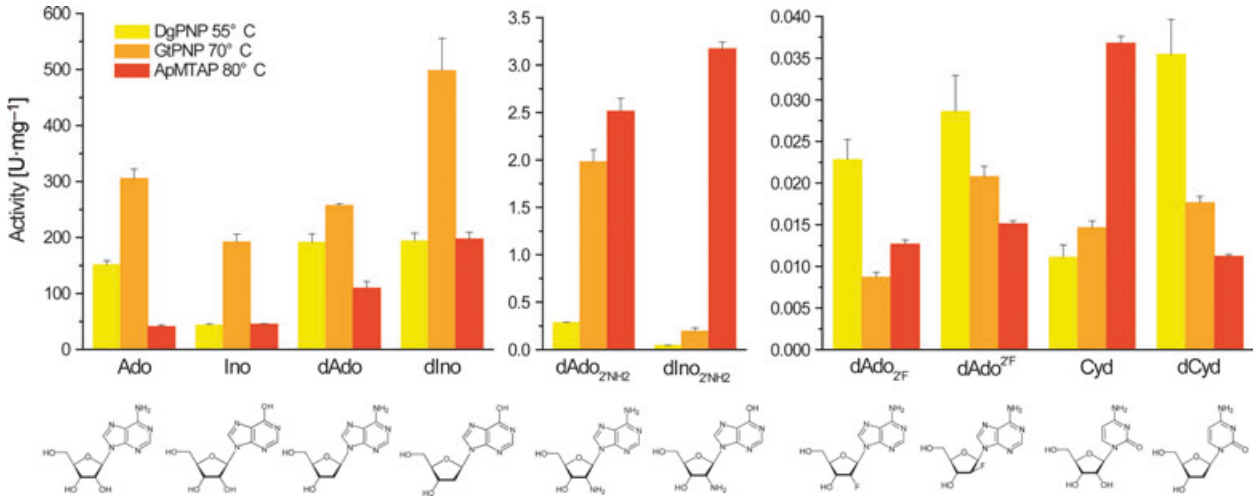


Fig. 4. Activity of the thermostable PNPs with natural and modified substrates. Reaction conditions: 1 mM substrate in 50 mM potassium phosphate buffer (pH 7) at 55 °C (DgPNP), 70 °C (GtPNP) and 80 °C (ApMTAP). Reactions were stopped by trichloroacetic acid or cold MeOH (dAdo and dIno) in the initial reaction phase.

Phosphorolysis of pyrimidine nucleosides

Araki *et al.* patented the method for producing cytidine compounds using recombinant *E. coli* PNP as biocatalyst and 2-deoxy-α-D–ribofuranose-1-phosphate and cytosine as substrates in 100 mM Tris/HCl buffer [36]. Recently, Stepchenko *et al.* described the phosphorolysis of 2′–deoxycytidine (dCyd), 2′–deoxyuridine (dUrd) and 2′–deoxythymidine (dThd) under the conditions of recombinant *E. coli* PNP in 80 mM phosphate buffer (pH 7.0) at 15 °C. It was shown that dCyd and dUrd are good substrates and ~ 50% was phosphorolyzed after 24 h, whereas substrate activity of dThd was unexpectedly lower [37]. These data prompted us to test a number of pyrimidine nucleosides as substrates for the PNPs under investigation. It was found that substrate efficiency of cytidine (Cyd) is increasing in the following order: DgPNP (12 $mU \cdot mg^{-1}$) < GtPNP (14 $mU \cdot mg^{-1}$) < ApMTAP (36 $mU \cdot mg^{-1}$). Interestingly, in the case of dCyd, the reverse decrease in substrate activity was observed, viz. DgPNP (34 $mU \cdot mg^{-1}$) > GtPNP (17 $mU \cdot mg^{-1}$) > ApMTAP (10 $mU \cdot mg^{-1}$). The activities of DgPNP, GtPNP and ApMTAP with Urd (0.97, 0.83 and 4.21 $mU \cdot mg^{-1}$, respectively) and thymidine (< 0.05, 0.26 and 6.83 $mU \cdot mg^{-1}$, respectively) were found to be rather low.

Kinetic parameters

The kinetic parameters of the thermostable PNPs are summarized in Table 2, along with other literature data. K_m and V_{max} were measured in 50 mM potassium phosphate buffer and typical Michaelis–Menten kinetics were observed. Note that TtPNP, EcPNP, SsMTAP and PfMTAP are hexamers, whereas TtPNPII is a monomer.

The results show that the enzymes have a higher affinity (low K_m values) towards adenosine than to inosine, e.g. 5–8 times higher in the cases of DgPNP and GtPNP. Similar remarkable differences in the substrate affinity were also described for PNPs from *T. thermophilus* [18] and PfMTAP [26]. The k_{cat} values of the nucleoside phosphorylases assessed for the natural substrates dIno and dAdo point to a very high catalytic activity of GtPNP (cf. the relevant data in Fig. 4), provoking reasonable interest to its use in biotechnological transformations.

Multiple sequence alignment

DgPNP, GtPNP and ApMTAP share a sequence identity with EcPNP of 54, 55 and 30%, whereas GtPNP shows 86% identity with GsPNP; ApMTAP shows 45% identity with SsMTAP.

The amino acid sequence alignment of 'high molecular mass PNPs' from different thermopiles with *E. coli* is shown in Fig. 5. The thermal characteristics and substrate specificities of these proteins were described above.

The alignment shows that the amino acid residues known to be involved in the active site of *E. coli* PNP and *S. solfataricus* MTAP [19,24,39] are generally conserved, with the exception of Met64, Ser90, Cys91, Ala156, Pro162 and Phe167 in EcPNP which have variations in the thermostable PNPs (indicated in boxes in Fig. 5).

The amino acid residues from 62 to 76 of EcPNP are considered to be highly conserved sequence motifs [39] involved in ribose binding [19]. In the thermozymes, especially in ApMTAP (Fig. 5), smaller residues such as Ile66*, Gly68, Ala71, Ala72, Val73* and Val74* replace Met64*, Ile66, Cys69, Ser70* and Ile71* of EcPNP, respectively (residues highlighted in pink of Fig. 5 are designated by * in the text). Moreover, in or close to the phosphate- and base-binding sites, similar phenomena can be observed: smaller residues in ApMTAP (or SsMTAP) Gly25*, Ala44*, Ala164*, Ala167, Ala172 and Ala173* replace Leu23*, Val42*, Leu158*, Ser161, Met166 and Phe167* of EcPNP, respectively. On the one hand, the smaller residues provide more space for the modified substrate [38], which may explain the relatively broad substrate spectrum of ApMTAP. On the other hand, the amino acid composition of thermophilic proteins is thought to be correlated with their thermostability [13]. Indeed, compared with EcPNP, ApMTAP contains more hydrophobic and small residues such as Val (10.4 versus 8.4%), Leu (12.3 versus 6.3%) and Gly (13.9 versus 9.6%). Furthermore, it is known that the substitution of Lys by Arg residues is advantageous at high temperatures because the Arg δ–guanidino moiety has less reactivity than the Lys amino group due to its high pK_a and resonance stabilization [13]. It can be observed that the Arg content increases from EcPNP (5.0%), GtPNP (6.0%), DgPNP (6.9%) to ApMTAP (8.2%), whereas the Lys content decreases from EcPNP (5.9%), GtPNP (3.4%), ApMTAP (1.6%) to DgPNP (0.8%).

Among the eight PNPs shown in Fig. 5, only TtPNPI [18] cannot accept Ado as a substrate, supposedly because of its residues substitutions in the base-binding sites: Glu156, Phe178 and Asn204 instead of Ala156, Val178 and Asp204 in EcPNP. Note that these residues are highly conserved in other PNPs (Fig. 5). The Asp204 (EcPNP) is considered as

Table 2. Kinetic parameters of PNPs on inosine and adenosine. NA, data not available.

Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})[a]	k_{cat}/K_m($s^{-1}\cdot\mu M^{-1}$)	V_{max} ($U\cdot mg^{-1}$)	Reaction conditions	Ref.
DgPNP	Inosine	975 ± 65	43 ± 1	0.043	85 ± 2	55 °C, 50 mM phosphate, pH 7.0	This study
	Adenosine	212 ± 18	91 ± 2	0.43	182 ± 4		
GtPNP	Inosine	1801 ± 288	248 ± 13	0.14	540 ± 28	70 °C, 50 mM phosphate, pH 7.0	This study
	Adenosine	240 ± 26	172 ± 6	0.72	375 ± 14		
ApMTAP	Inosine	137 ± 22	23 ± 1	0.17	51 ± 2	80 °C, 50 mM phosphate, pH 7.0	This study
	Adenosine	97 ± 14	20 ± 1	0.21	44 ± 2		
SsMTAP	Inosine	84 ± 4	22 ± 1	0.26	NA	70 °C,100 mM phosphate, pH 7.4	[24]
	Adenosine	25 ± 1	7 ± 1	0.29	NA		
PfMTAP	Inosine	963	1.6	0.0016	NA	80 °C,100 mM phosphate, pH 7.4	[26]
	Adenosine	109	3.8	0.035	NA		
TtPNPI	Inosine	2080 ± 20	29 ± 1	0.014	69 ± 1	80 °C, 50 mM phosphate, pH 7.0	[18]
TtPNPII	Adenosine	320 ± 20	440 ± 3	1.38	880 ± 6		
EcPNP	Inosine	96 ± 2	40 ± 1	0.72	10 [b]	25 °C, 50 mM phosphate, pH 7.4	[38]
	Adenosine	46 ± 7	33 ± 4	0.42	8 [b]		

[a] Calculations of k_{cat} are based on the subunit molecular mass of the enzymes. [b] Data not reported but kindly provided by Bennett *et al.* [38].

the key base-binding amino acid residue by the protonation of the purine base at N7 accompanied by hydrogen bonding with the C6 amino group [19]. According to Mikleušević *et al.* [40], mutation of Arg204 to Asn in EcPNP results in a 1000–fold decrease in the enzymatic activity, which further provides evidence for the role of Asp204 in the PNP specificity of Ado. Asp204 (corresponding to Asp223 in DgPNP, Asp 203 in GtPNP and Asp210 in ApMTAP) was conserved in the PNPs studied here and as a consequence they can well accept Ado as substrate. However, to explain the difference in the reaction rate and substrate preference between the studied PNPs and EcPNP (see Table 2), profound structure studies of the catalytic site are needed.

Synthesis of fluorinated purine nucleosides

The aforementioned results of this study and data from previous studies [17] prompted us to investigate the synthesis of adenine 2′-deoxyfluoro-*ribo*- and -*arabino*-nucleosides, $dAdo_{2'F}$ (**17**) and $dAdo^{2'F}$ (**18**), respectively, using adenine as an acceptor and the respective 2′-fluorinated uracil nucleosides, $dUrd_{2'F}$ (**13**) and $dUrd^{2'F}$ (**14**), as donors of the pentofuranose residues (Scheme 3). Preliminary experiments showed the formation of: (a) $dAdo_{2'F}$ using TtPyNP and DgPNP as biocatalysts (55 °C, 24 h) with a yield of 23%, and (b) $dAdo^{2'F}$ employing a TtPyNP and ApMTAP combination of enzymes (80 °C, 24 h) with a yield of 24%. These data point to the possibility of developing a practical synthesis of both 2′-fluorinated adenine nucleosides, although an optimization of the reaction conditions is necessary to increase the yield.

Pyrimidine and purine nucleosides modified at the C2′ carbon atom of the pentofuranose ring are of great importance as constituents of artificial oligonucleotides of medicinal potential [41,42]. Among them, 2′-amino-2′-deoxy- and 2′-deoxy-2′-fluoro-β-D–*ribo*-nucleosides attracted much attention during the last two decades and a number of oligonucleotides containing these modified nucleosides at diverse position of the oligonucleotide chain have been synthesized and their properties investigated [43–45]. More recently, 2′-deoxy-2′-fluoro-β-D–*arabino*-nucleosides have been shown to be of great importance for antisense and low molecular mass oligonucleotides as potential anticancer drugs [6,46,47]. Unfortunately, chemical synthesis of pyrimidine and especially purine 2′-deoxyfluoro-*ribo*- and -*arabino*-nucleosides, as well as their 2′-aminodeoxy counterparts suffers from many drawbacks despite of enormous progress achieved in this field of research. The search for enzymatic transformations that can efficiently replace some of the chemical steps in the production of these nucleosides using the principles of 'green chemistry' is therefore of great importance.

In our previous work, we described the recombinant expression and biocatalytic characterization of two thermostable PyNPs, isolated from *G. thermoglucosidasius* (GtPyNP) and *T. thermophilus* (TtPyNP) [17]. It was found that GtPyNP displays very low phosphorolytic activity towards $dUrd_{2'F}$ and $dUrd^{2'F}$, whereas TtPyNP shows rather satisfactory activity with both 2′-deoxyfluoro-*ribo*- and -*arabino*-nucleosides. This observation pointed to the possibility of employing TtPyNP as a biocatalyst in the transglycosylation reaction with adenine as the pentofuranosyl acceptor and

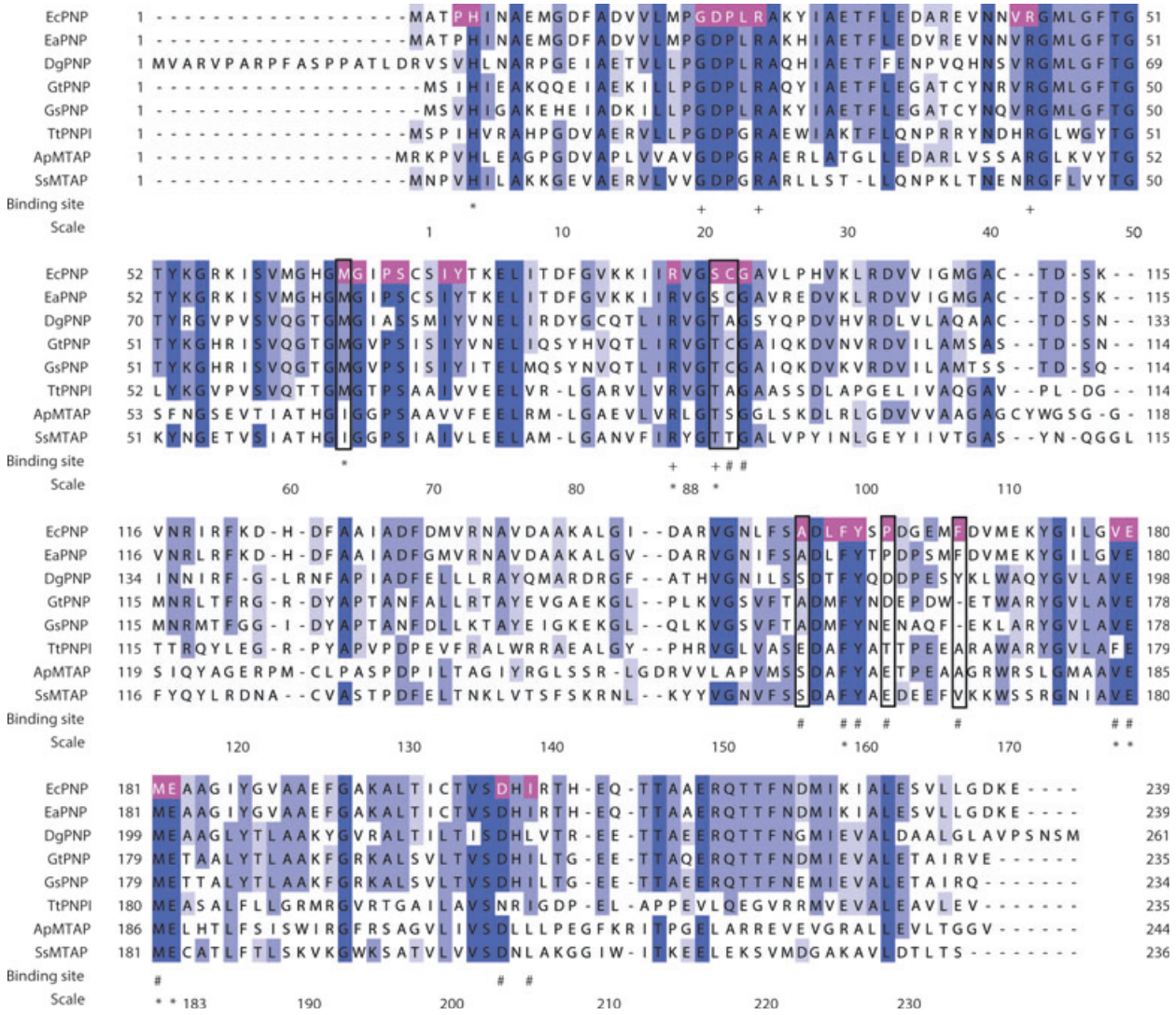
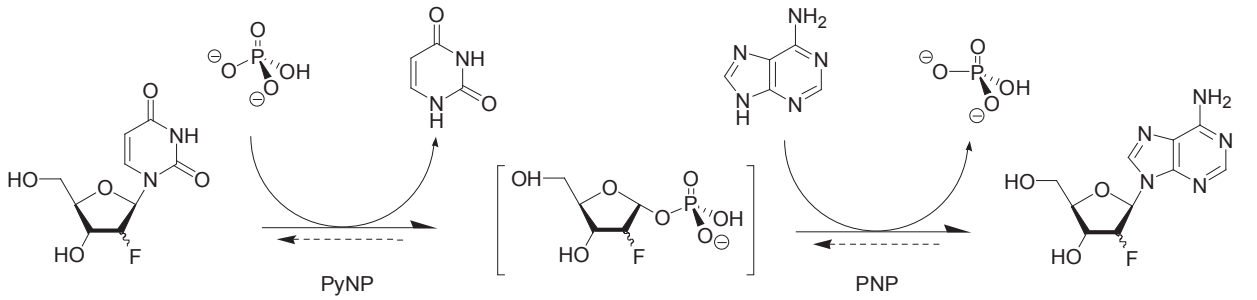


Fig. 5. Multiple sequence alignment of high molecular mass PNP or MTAP from *E. coli* (EcPNP), *Enterobacter aerogenes* (EaPNP), *D. geothermalis* (DgPNP), *G. thermoglucosidasius* (GtPNP), *G. stearothermophilus* (GsPNP), *T. thermophilus* (TtPNPI), *A. pernix* (ApMTAP) and *S. solfataricus* (SsMTAP), respectively. The scale below is based on the EcPNP sequence reported in the literature [19,24]. The shading represents the degree of sequence identity. Residues of a representative sequence involved in the active site or in intersubunit contacts Bzowska *et al.* [19] and Pugmire *et al.* [39] are highlighted in pink. The ribose (*), phosphate (+) and base (#) binding sites according to Bzowska *et al.* [19] and Cacciapuoti *et al.* [24] are indicated below the sequences.

$dUrd_{2'F}$ or $dUrd^{2'F}$ as donors of the corresponding *ribo*- and *arabino*-pentofuranose.

Here, we studied this reaction testing DgPNP, GtPNP and ApMTAP as biocatalysts for the coupling of the intermediates 2′-deoxyfluoro-*ribo*- and *arabino*-1-phosphates with adenine. It was found that DgPNP catalyzes the formation of the desired $dAdo_{2'F}$ under very mild conditions (55 °C, 24 h) with a yield of 24%. It is noteworthy that the undesired chemical transformation of $dUrd_{2'F}$ into O^2,2′–anhydro-1-(β-D–arabinofuranosyl)uracil was very low (2%) at 55 °C [17]. The use of TtPyNP and ApMTAP as a combination of enzymes, and $dUrd^{2'F}$ and adenine as substrates (80 °C, 24 h) resulted in the formation of $dAdo^{2'F}$ with a yield of 24%; TtPyNP and GtPNP as a combination of enzymes for the same synthesis resulted in the formation of $dAdo^{2'F}$ with a yield of 14%. It is obvious that these results are very promising from the viewpoint of their application in the development of practical syntheses of $dAdo_{2'F}$ and $dAdo^{2'F}$.



Scheme 3. Synthesis of fluorinated purine nucleosides by the coupled reactions of PyNP and PNP.

Phosphorolysis of purine nucleosides catalyzed by GtPyNP and TtPyNP

During the course of the abovementioned studies, we observed that the thermostable PyNPs, GtPyNP and TtPyNP, phosphorolytically cleaved not only pyrimidine nucleosides [17], but also various purine nucleosides. To investigate this phenomenon further, we determined the temperature dependence of PyNP activity with inosine in order to exclude the possibility that the PNP activity comes from *E. coli* PNP. It was found that temperature dependence of GtPyNP and TtPyNP activity with inosine (Fig. 6) is similar to that obtained in experiments with one of the natural substrates, uridine (see Fig. 4A in Ref. [17]). Hence, we assume that the PNP activity is inherent to the PyNPs and not the result of contamination with *E. coli* PNP.

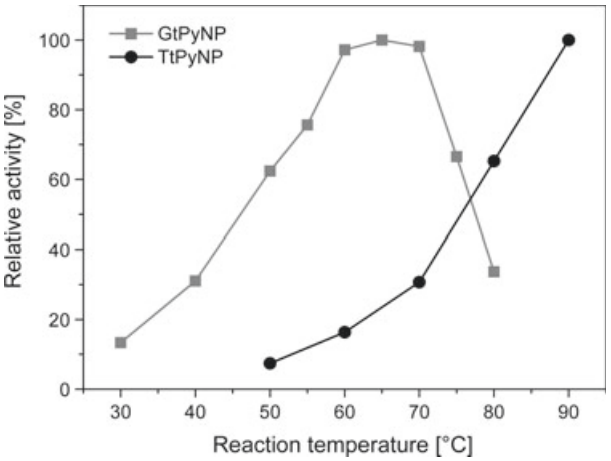


Fig. 6. Temperature dependence of the activity of GtPyNP and TtPyNP on inosine. The highest reaction rate observed for each enzyme was set to 100%. Substrate: 1 mM inosine in 50 mM potassium phosphate buffer (pH 7).

Table 3. Activity of GtPyNP and TtPyNP with natural and non-natural nucleosides. Reaction conditions: 1 mM substrate in 50 mM phosphate buffer at 60 °C (GtPyNP) or 80 °C (TtPyNP).

	Activity ($U \cdot mg^{-1}$)	
Substrate	GtPyNP 60 °C	TtPyNP 80 °C
Urd	51.5	153.7
Thd	22.4	295.2
Ado	0.458	0.809
Ino	0.308	0.900
$dAdo_{2'NH2}$	0.021	0.065
$dIno_{2'NH2}$	0.012	0.067
$dUrd_{2'F}$	0.002[a]	0.052[a]
$dUrd^{2'F}$	ND[a,b]	0.005[a]
Cyd	ND[b]	0.001

[a] 10 mM phosphate buffer. [b] Not observed.

Furthermore, the activities of GtPyNP (at 60 °C) and TtPyNP (at 80 °C) with the other purine nucleosides were examined and the results are summarized in Table 3. It is obvious that both PyNPs display considerable activities with Ado and Ino. In fact, TtPyNP accepts the purine significantly better than 2′–fluorinated pyrimidine nucleosides. By contrast, the activity with the natural pyrimidine nucleoside cytidine was extremely low (TtPyNP) or was not detectable (GtPyNP).

Discussion

In this study, we prepared three recombinant PNPs from thermophiles and studied their properties and substrate specificity for natural substrates (Ado, dAdo, Ino, dIno) and three sets of analogues, viz., (a) $dAdo_{2'NH2}$ and $dIno_{2'NH2}$, (b) $dAdo_{2'F}$ and $dAdo^{2'F}$, (c) Cyd and dCyd. The reasons for selection of the analogues are as follows: the first have closest structural relation of the relevant natural purine nucleosides; the diastereomeric deoxyfluoro adenine nucleosides also

structurally mimic dAdo. However, it is well known that the replacement of a hydrogen atom or hydroxyl group in the sugar moiety of nucleosides gives rise to derivatives with unexpected physico-chemical and biological properties. Finally, pyrimidine nucleosides are not typical substrates of PNP, however, Cyd and dCyd mimic the pyrimidine fragment of adenine base of N^3–(β-D–ribofuranosyl)adenine that was shown to be the substrate of PNP [48–50] and a good substrate activity of dCyd and 2′-deoxyuridine for *E. coli* PNP was also described [36,37].

The role of the vicinal C2′ and C3′ hydroxyl groups of natural purine nucleosides in the binding site of PNPs has been discussed in a number of publications [38,51,52]. The PNPs studied here and *E. coli* PNP contain the same cluster of amino acids (Val–Glu–Met –Glu). In this cluster, the last residue Glu181 of *E. coli* PNP was suggested to play an important role in the binding of ribose *via* the vicinal C2′ and C3′ hydroxyl groups and in facilitating substrate activation by flattering the pentofuranose ring [38]. In a similar way, one can suggest that the relevant Glu residues of the enzymes under study [DgPNP (Glu200), GtPNP (Glu180) and ApMTAP (Glu187); (Fig. 5)] are involved in the ribose binding site. It was previously shown that *E. coli* PNP accepts the natural purine *ribo*- and 2′-*deoxyribo*-nucleosides with similar efficiency. The results of this study also showed a similar trend (Fig. 4), indicating that the C2′ hydroxyl groups of Ado and Ino are not essential for the enzymatic reaction. Note that there is a lot of theoretical and experimental data convincingly demonstrating that natural *ribo*- and 2′-*deoxyribo*-nucleosides have rather high conformational flexibility, viz., S ↔ N conformation of the pentofuranose ring and base rotation around the glycosidic bond. These conformational peculiarities allow PNPs adopting the required spatial structure of a natural nucleoside substrate in the catalytic site, which is necessary to form the productive enzyme-substrate complex.

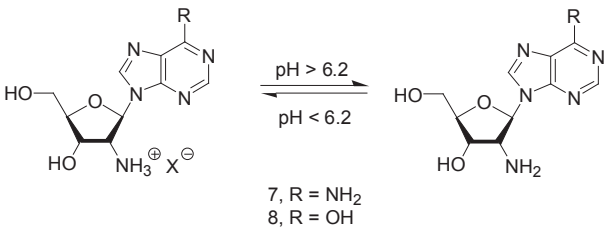
Compounds of the first and second sets are functionally competent, i.e. contain all the functionality of inherent natural substrate, purine nucleosides and, *a priori*, one can expect their substrate activity for PNPs to be similar to those of natural substrates. Indeed, the amino function of $dAdo_{2'NH2}$ and $dIno_{2'NH2}$ is isosteric for the hydroxyl group and can functionally replace the hydroxyl group. The only important peculiarity of the amino function vs hydroxyl group consists in that the former can exist in the C2′-ammonium cation ↔ C2′-amino equilibrium implying the pH dependence of the substrate activity, on the one hand (Scheme 4), and the ionic interaction of the phosphate anion and ammonium cation impeding the correct spatial arrangement of the productive nucleoside/phosphate/enzyme complex, on the other.

The observed dramatic decrease in substrate activity by going from natural substrates to $dAdo_{2'NH2}$ (**7**) and $dIno_{2'NH2}$ (**8**) (2–3 orders of magnitude) prompted us to: (a) test the substrate activity of both aminodeoxy nucleosides at different pH values of the reaction mixtures (Table S3, Fig. S1), and (b) analyze the C2′-ammonium cation and C2′-amino forms of the former by the restricted Hartree–Fock (RHF) *ab initio* method using a basis set of STO-3G (HYPER-CHEM, 8.1 release) and the PM3 geometry optimization as starting approximation for the *ab initio* calculations compared to similar data for Ado (Table S4).

Preliminary experiments on the pH dependence on the substrate properties gave a rather variable picture of activities (Fig. S1) and pointed to the need to perform more detailed studies to obtain meaningful results. For example, a plot of k_{cat}/K_m as a function of pH is needed to reflect the essential ionizing groups of the free enzyme that play a role in both substrate binding and catalytic processing [53].

Analysis of the amino and ammonium forms of $dAdo_{2'NH2}$ by the *ab initio* method revealed the essential differences in the base orientation about the glycosyl bond as well as the ribofuranose rings of two species, on the one hand, and some difference of both forms from the natural substrate, on the other hand. Conformational analysis of $dAdo_{2'NH2}$ (**7**) in water by NMR spectroscopy clearly pointed to its high S ↔ N and *anti* ↔ *syn* conformational flexibility [54]. Taking into account that both functionally competent forms are in equilibrium under optimal pH values for PNP functioning, an enzyme can accept the most suitable conformer for binding. On the whole, these considerations allow us to suggest the involvement of the C2′-ammonium cation in an interaction with phosphate anion as a likely reason for the very low substrate activity of $dAdo_{2'NH2}$ (**7**) and $dIno_{2'NH2}$ (**8**).

In the case of diastereomeric adenine deoxyfluoro nucleosides $dAdo_{2'F}$ (**9**) and $dAdo^{2'F}$ (**10**), the most



Scheme 4. Equilibrium of C2′-ammonium cation and C2′-amino at different pH environment.

unexpected finding is that both functionally competent analogues show very similar extremely low (3–4 orders of magnitude compared with that of Ado) substrate activity for the all investigated enzymes. The H and F atoms have similar van der Waals' radii and thus do not create any steric hindrances to the substrate binding in the catalytic center of the investigated enzymes. However, a fluorine atom, as an isostere for the hydrogen atom [55], differs dramatically from its isostere owing to the strongest electronegative character exerting diverse influences on electronic properties of proximal atoms as well as on the stereochemistry of the pentofuranose ring.

An *ab initio* analysis of $dAdo_{2'F}$ and $dAdo^{2'F}$ and comparison of the results obtained with those of X–ray crystallographic data for Ado and $dAdo_{2'F}$ (Table S4), as well as the data of the conformational NMR analysis of $dAdo_{2'F}$ [56] and $dAdo^{2'F}$ [57] unexpectedly revealed a rather broad stereochemical and electronic similarity of the 2′-deoxyfluoro-*ribo*- and -*arabino*-nucleosides that is in harmony with the activity data (Fig. 4). The only reasonable explanation of the observed very low substrate activity of $dAdo_{2'F}$ and $dAdo^{2'F}$ may be connected to the structural peculiarities of the productive complex giving rise to the glycoside bond cleavage. Bennet *et al.* suggested flattering of the pentofuranose ring of the substrate (the C4′–O4′–C1′–C2′ torsion angle strives to 0°) as one of the most important features of the reaction [38]. As distinct from the Ado/ara-adenine (Ade) pair, the C2′ fluorine atom of $dAdo_{2'F}$ and $dAdo^{2'F}$ in all likelihood exerts a much higher energy barrier than that of Ado and dAdo for flattering of the pentofuranose ring, which is unexpectedly similar to and independent of the fluorine *ribo*- or *arabino*- configuration.

Conclusion

Here, we described the preparation of three PNPs from the thermophilic microorganisms *D. geothermalis*, *G. thermoglucosidasius* and *A. pernix*, aiming at the search for new efficient biocatalysts for the synthesis of modified nucleosides. The recombinant expression in *E. coli* resulted in a high yield of biologically active enzymes. The PNPs were found to be stable at their optimal temperatures (DgPNP: optimal activity at 55 °C retaining thereupon 80% for 8 h; GtPNP: optimal activity at 70 °C without losing activity for 8 h; ApMTAP: activity exponentially rising up to 99 °C). Natural substrates (Ado, Ino and their 2′-deoxy counterparts) are effectively phosphorolyzed by all three investigated enzymes (activity level within 50–500U·mg^{-1}). The activity of the enzymes for $dAdo_{2'NH2}$ and $dIno_{2'NH2}$ was found to be approximately two orders of magnitude lower (within 0.01–3.25 U·mg^{-1}) than their activity for natural substrates. ApMTAP showed the highest level of enzymatic activity for $dAdo_{2'NH2}$ and $dIno_{2'NH2}$. A similar further decrease in substrate activity was found for $dAdo_{2'F}$ and $dAdo^{2'F}$ (within 0.01–0.03 U·mg^{-1}). DgPNP showed the highest activity for this type of substrate among the enzymes tested here. It should be stressed that the substrate activities of $dAdo_{2'F}$ and $dAdo^{2'F}$ are unexpectedly similar and do not depend on the fluorine atom configuration. Interestingly, cytidine was phosphorolyzed by the all studied PNPs with an efficiency similar to that of adenine 2′-deoxyfluoro nucleosides. Furthermore, uridine and thymidine were also phosphorolyzed by the studied PNPs (except for DgPNP, which cannot accept thymidine as substrate), but the activity was 6–30 times lower than that for cytidine.

It was found that the previously described thermostable pyrimidine nucleoside phosphorylase TtPyNP [17] and the present described DgPNP and ApMTAP can be applied for the transglycosylation reaction of adenine in the presence of $dUrd_{2'F}$ and $dUrd^{2'F}$ with the formation of the corresponding adenine nucleosides $dAdo_{2'F}$ and $dAdo^{2'F}$ at a yield of ~ 24%.

Table 4. Oligonucleotide primers and GenBank number of PNPs in this study.

Enzyme	Primer sequence (5′–3′)[a]	GenBank Accession number
DgPNP	F:TACTAG**GGATCC**GTGGCGCGTGTACCGG	ABF45792
	R:CAGCAT**AAGCTT**TCACATGCTGTTGGAAGG	
GtPNP	F:ACTAG**GGATCC**TTGAGCATCCATATCGAAG	AEH47728
	R:AGCAT**AAGCTT**TTATTCTACGCGAATCGCC	
ApMTAP	F:ACTAG**GGATCC**AGGAAGCCGGTTCACCTCG	NP_147653
	R:AGCAT**AAGCTT**CTAGACTCCTCCTGTGAGG	

[a] Restriction sites for *Bam*HI (**GGATCC**) and *Hind*III (**AAGCTT**) are indicated in bold and the insert-specific sequences are underlined. F, forward primer; R, reverse primer.

The investigation of the underlying enzymatic mechanisms of the prepared thermostable PNPs and their employment for the synthesis of purine nucleosides of biological interest are in progress.

Materials and methods

Bacterial strains, cloning, expression and purification

Escherichia coli TOP10 (Invitrogen, Carlsbad, CA, USA) and BL21 (Novagen, Darmstadt, Germany) strains were used for gene cloning and protein expression, respectively.

The genomic DNA of *T. thermophilus* HB27 and *G. thermoglucosidasius* 11955 were isolated as previously reported [17]. The genomic DNA of *D. geothermalis* and *A. pernix* were isolated by Marco Casteleijn (Bioprocess Engineering Laboratory of the University of Oulu, Finland). The target gene sequences were amplified from genomic DNA by PCR using *Pfu* DNA polymerase (Fermentas, Vilnius, Lithuania). The used primer pairs are listed in Table 4.

The PNP encoding genes were cloned via *Bam*HI/*Hin*dIII digestion (FastDigest restriction endonucleases; Fermentas) and subsequent ligation (T4 DNA Ligase; Roche, Mannheim, Germany) into a modified pCTUT7 vector [17]. In brief, this vector includes an IPTG inducible *lac* promoter and a hexahistidine coding sequence connected to the 5′–end of the target gene.

DgPNP and ApMTAP were expressed in *E. coli* BL21 in EnPresso® medium (BioSilta, Oulu, Finland) by the use of the Enbase® technology [58] using Ultra Yield Flasks™ and AirOTop™ seals (both from BioSilta, Oulu, Finland). The expression of DgPNP and ApMTAP was induced after overnight cultivation at 37 or 30 °C, respectively, by the addition of 20 μM IPTG. GtPNP was expressed in *E. coli* BL21 cells in Terrific Broth (TB) medium [59] by the induction of 100 μM IPTG to the culture growing at 37 °C for 2.5 h. Cells were harvested by centrifugation (16 000 ***g***, 5 min, 4 °C) after either 24 h (for EnPresso® medium) or 3.5 h (for TB medium) after induction, respectively.

Cell disruption and protein purification were carried out as described previously [17]. Briefly, cells were broken by ultrasonic treatment (UP200S sonicator; Hielscher Ultrasonics GmbH, Teltow, Germany), then the lysate was heat-treated at 50 °C (DgPNP), 65 °C (GtPNP) or 85 °C (ApMTAP) for 15 min, and after centrifugation (16 000 ***g***, 20 min, 4 °C) the supernatant containing the target protein was further purified by metal ion-affinity chromatography (Ni-NTA Superflow cartridge; Qiagen, Hilden, Germany) and the HiPrep 26/10 desalting column on an Äkta FPLC system (GE Healthcare, Munich, Germany).

The protein purity was checked by SDS/PAGE according to a standard protocol [59]. To analyze whether the disulfide bridge exists in the enzymes, the purified enzymes (0.5 $mg \cdot mL^{-1}$) were treated with or without the reducing agent, i.e. 0.1 M dithiothreitol, in the incubation step of SDS/PAGE analysis. The molecular mass marker (Sm0431) was purchased from Fermentas. The purified protein concentration was determined by using a NanoDrop1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, ME, USA) [60] and applying $E_{280\ nm}^{1\%} = 6.3$ (for DgPNP), or 8.8 (for GtPNP and ApMTAP), respectively. The absorption coefficients were theoretically calculated from the amino acid sequence (VECTOR NTI software; Invitrogen).

Enzyme activity assay

All the reactions were performed within 200–1000 μL scale in Eppendorf tubes, which were incubated in a thermomixer (Biozym, Hessisch Oldendorf, Germany) at 300 rpm. The standard activity assay (phosphorolysis) was carried out in potassium phosphate buffer (50 mM; pH 7.0) containing 1 mM purine nucleoside. After 2 min preheating at the corresponding temperature, a certain amount of purified enzyme (~ 0.5 $\mu g \cdot mL^{-1}$ for natural substrates; 0.1 $mg \cdot mL^{-1}$ for modified substrates) was added into the mixture and the reaction was stopped (addition of 1/2 vol of 10% trichloroacetic acid or methanol) after a defined time interval so that < 10% of the substrate was converted to the product. Under these conditions, the reaction rate was linear as a function of time and enzyme concentration. After centrifugation (20 000 ***g***, 15 °C, 25 min) the samples were stored at −20 °C for following analyses. Negative controls (without enzyme addition) were performed in parallel. The reaction mixtures were analyzed by HPLC as described in [17]. The substrate conversion was calculated from the HPLC chromatograms:

$$\text{Conversion}(\%) = \frac{[\text{product}]}{[\text{product}] + [\text{substrate}]} \times 100\%$$

One unit (U) of enzyme activity was defined as the amount of the enzyme catalyzing the conversion of 1 μmol of substrate per minute under the respective assay conditions.

Temperature optimum and thermal stability measurements

In order to determine temperature optima of the enzymes, the reaction mixture (1 mM inosine in 50 mM potassium phosphate buffer, pH 7.0) was preheated for 2 min at different temperatures (from 30 to 99 °C), suitably diluted enzyme solution was added, and the reaction was stopped by the addition of trichloroacetic acid after 3 min.

Thermal stability was analyzed by incubating purified PNP aliquots (13–55 $\mu g \cdot mL^{-1}$) in 50 mM phosphate buffer (pH 7.0) on a PCR machine (Eppendorf, Hamburg, Germany) at the respective temperatures. After defined time

intervals, tubes were withdrawn and cooled on ice. The residual activity of the incubated enzyme was determined at 50 °C for DgPNP, 60 °C for GtPNP and 80 °C for ApMTAP using inosine as a substrate under standard assay conditions. In order to determine the half-life, the residual activity data over time were fitted to the exponential decay equation $a = a_0 \cdot e^{(-\lambda \cdot t)}$ (SIGMAPLOT 11.0; Systat Software GmbH, Erkrath, Germany), where λ is the first-order deactivation coefficient, a is the residual activity and a_0 is the original activity before treatment. Finally, the half-life was obtained by $t_{1/2} = \ln 2 \cdot \lambda^{-1}$.

Kinetic parameters determination

Reactions were performed in triplicate at least to obtain the initial velocity at a fixed phosphate concentration (50 mM, pH 7.0) and the substrate concentration varied over a 20-fold range (spanning 0.25–5 times the Michaelis–Menten constant) [61]. Reaction temperatures were: 55 °C for DgPNP, 70 °C for GtPNP and 80 °C for ApMTAP. K_m and V_{max} were determined by nonlinear regression based on the Michaelis–Menten equation (applied in SIGMAPLOT 11.0).

Multiple sequence alignment

Protein identities were assessed with the protein basic local alignment tool (BLAST) of the NCBI web server [62]. The multiple alignments were constructed using Expresso mode of T-Coffee server [63], which implemented an automated identification of suitable structural templates via a BLAST search against the PDB database [64]. The corresponding UniProt Protein accession numbers are P0ABP8 (EcPNP), G0E416 (EaPNP), Q1IY92 (DgPNP), F8CYG4 (GtPNP), P77835 (GsPNP), Q72IR2 (TtPNPI), Q9YDC0 (ApMTAP) and P50389 (SsMTAP). The figure of alignment was created using Jalview2 [47].

Synthesis of fluorinated purine nucleosides

The syntheses of adenine 2′–deoxyfluoro-*ribo*- and -*arabino*-nucleosides ($dAdo_{2'F}$ and $dAdo^{2'F}$, respectively) were performed in phosphate buffer (2 mM, pH 6.5) containing 2 mM $dUrd_{2'F}$ or $dUrd^{2'F}$ and 1 mM Ade. The enzyme loading was 0.1 mg·mL^{-1} for each nucleoside phosphorylase, viz., GtPyNP or TtPyNP and PNP. The reaction was conducted at defined temperatures for up to 24 h. The synthesized nucleosides, $dAdo_{2'F}$ and $dAdo^{2'F}$ were identified by comparison of their retention times and UV spectra with those of authentic samples by HPLC (Fig. S2). The calculated yields are based on the amount of adenine applied into the reaction.

Preparation of thermostable PyNP, isolated from the thermophilic microorganisms *Geobacillus thermoglucosidasius* (GtPyNP) and *Thermus thermophilus* (TtPyNP) was described in [17].

Nucleosides and bases

Ade, Ado and Ino were purchased from Carl Roth (Karlsruhe, Germany); hypoxanthine, dAdo, dIno, Cyt and dCyt were purchased from Sigma-Aldrich (Steinheim, Germany); $dAdo^{2'F}$ was purchased from Metkinen Chemistry (Kuusisto, Finland); $dAdo_{2'NH2}$, $dIno_{2'NH2}$ and $dAdo_{2'F}$ were kindly provided by Prof. Alex Azhayev (Metkinen Chemistry).

Acknowledgements

We sincerely thank Dr Marco G. Casteleijn (Bioprocess Engineering Laboratory, University of Oulu, Finland) for providing us with the genomic DNA of *D. geothermalis* and *A. pernix*. We further thank Prof. Alex Azhayev (Metkinen Chemistry) for the donation of the substrates of 1-(2-deoxy-2-fluoro-β-D–arabinofuranosyl)uracil, 2′–amino-2′–deoxyadenosine and 2′–amino-2′–deoxyinosine. We are thankful to the Alexander von Humboldt Foundation (Bonn–Bad–Godesberg, Germany) for computer and program facilities used here. We are also indebted to Prof. Steven E. Ealick (Baker Laboratory, Cornell University) and Dr William Bill Parker (Southern Research Institute, Birmingham) for kindly providing the V_{max} data for *E. coli* PNP. This work is part of the Cluster of Excellence 'Unifying Concepts in Catalysis' coordinated by the Technische Universität Berlin. Financial support by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the German Initiative for Excellence is gratefully acknowledged (EXC 314).

References

1 De Clercq E (2009) The history of antiretrovirals: key discoveries over the past 25 years. *Rev Med Virol* **19**, 287–299.
2 Flexner C (2007) HIV drug development: the next 25 years. *Nat Rev Drug Discov* **6**, 959–966.
3 Bonate PL, Arthaud L, Cantrell WR Jr, Stephenson K, Secrist JA III & Weitman S (2006) Discovery and development of clofarabine: a nucleoside analogue for treating cancer. *Nat Rev Drug Discov* **5**, 855–863.
4 Watts JK & Corey DR (2012) Silencing disease genes in the laboratory and the clinic. *J Pathol* **226**, 365–379.

5 Manoharan M, Akinc A, Pandey RK, Qin J, Hadwiger P, John M, Mills K, Charisse K, Maier MA, Nechev L *et al*. (2011) Unique gene-silencing and structural properties of 2′–fluoro-modified siRNAs. *Angew Chem Int Edit* **50**, 2284–2288.
6 Watts JK & Damha MJ (2008) 2′F-arabinonucleic acids (2′F–ANA) – history, properties, and new frontiers. *Can J Chem* **86**, 641–656.
7 Ubiali D, Serra CD, Serra I, Morelli CF, Terreni M, Albertini AM, Manitto P & Speranza G (2012) Production, characterization and synthetic application of a purine nucleoside phosphorylase from *Aeromonas hydrophila*. *Adv Synth Catal* **354**, 96–104.
8 Mikhailopulo IA & Miroshnikov AI (2010) New trends in nucleoside biotechnology. *Acta Nat* **2**, 36–59.
9 Taran S, Verevkina K, Feofanov S & Miroshnikov A (2009) Enzymatic transglycosylation of natural and modified nucleosides by immobilized thermostable nucleoside phosphorylases from *Geobacillus stearothermophilus*. *Russ J Bioorg Chem* **35**, 739–745.
10 Hori N, Watanabe M, Sunagawa K, Uehara K & Mikami Y (1991) Production of 5–methyluridine by immobilized thermostable purine nucleoside phosphorylase and pyrimidine nucleoside phosphorylase from *Bacillus stearothermophilus* JTS 859. *J Biotechnol* **17**, 121–131.
11 Krenitsky TA, Koszalka GW & Tuttle JV (1981) Purine nucleoside synthesis, an efficient method employing nucleoside phosphorylases. *Biochemistry* **20**, 3615–3621.
12 Tuttle JV, Tisdale M & Krenitsky TA (1993) Purine 2′-deoxy-2′-fluororibosides as antiinfluenza virus agents. *J Med Chem* **36**, 119–125.
13 Vieille C & Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**, 1–43.
14 Reuter S, Rusborg Nygaard A & Zimmermann W (1999) β–Galactooligosaccharide synthesis with β-galactosidases from *Sulfolobus solfataricus*, *Aspergillus oryzae*, and *Escherichia coli*. *Enzyme Microb Technol* **25**, 509–516.
15 Kujo C & Ohshima T (1998) Enzymological characteristics of the hyperthermostable NAD-dependent glutamate dehydrogenase from the archaeon *Pyrobaculum islandicum* and effects of denaturants and organic solvents. *Appl Environ Microbiol* **64**, 2152–2157.
16 Hei DJ & Clark DS (1994) Pressure stabilization of proteins from extreme thermophiles. *Appl Environ Microbiol* **60**, 932–939.
17 Szeker K, Zhou X, Schwab T, Casanueva A, Cowan D, Mikhailopulo IA & Neubauer P (2012) Comparative investigations on thermostable pyrimidine nucleoside phosphorylases from *Geobacillus thermoglucosidasius* and *Thermus thermophilus*. *J Mol Catal B Enzym* **84**, 27–34.
18 Almendros M, Berenguer J & Sinisterra J-V (2012) *Thermus thermophilus* nucleoside phosphorylases active in the synthesis of nucleoside analogues. *Appl Environ Microbiol* **78**, 3128–3135.
19 Bzowska A, Kulikowska E & Shugar D (2000) Purine nucleoside phosphorylases: properties, functions, and clinical aspects. *Pharmacol Ther* **88**, 349–425.
20 Hamamoto T, Noguchi T & Midorikawa Y (1997) Cloning of purine nucleoside phosphorylase II gene from *Bacillus stearothermophilus* TH 6–2 and characterization of its gene product. *Biosci Biotechnol Biochem* **61**, 276–280.
21 Utagawa T, Morisawa H, Yamanaka S, Yamazaki A, Yoshinaga F & Hirose Y (1985) Properties of nucleoside phosphorylase from *Enterobacter aerogenes*. *Agric Biol Chem* **49**, 3239–3246.
22 Wei XK, Ding QB, Zhang L, Guo YL, Ou L & Wang CL (2008) Induction of nucleoside phosphorylase in *Enterobacter aerogenes* and enzymatic synthesis of adenine arabinoside. *J Zhejiang Univ Sci B* **9**, 520–526.
23 Zhu S, Ren L, Wang J, Zheng G & Tang P (2012) Two-step efficient synthesis of 5–methyluridine via two thermostable nucleoside phosphorylase from *Aeropyrum pernix*. *Bioorg Med Chem Lett* **22**, 2102–2104.
24 Cacciapuoti G, Forte S, Moretti MA, Brio A, Zappia V & Porcelli M (2005) A novel hyperthermostable 5′–deoxy-5′–methylthioadenosine phosphorylase from the archaeon *Sulfolobus solfataricus*. *FEBS J* **272**, 1886–1899.
25 Cacciapuoti G, Porcelli M, Bertoldo C, De Rosa M & Zappia V (1994) Purification and characterization of extremely thermophilic and thermostable 5′–methyl-thioadenosine phosphorylase from the archaeon *Sulfolobus solfataricus*. Purine nucleoside phosphorylase activity and evidence for intersubunit disulfide bonds. *J Biol Chem* **269**, 24762–24769.
26 Cacciapuoti G, Bertoldo C, Brio A, Zappia V & Porcelli M (2003) Purification and characterization of 5′–methylthioadenosine phosphorylase from the hyperthermophilic archaeon *Pyrococcus furiosus* – substrate specificity and primary structure analysis. *Extremophiles* **7**, 159–168.
27 Szeker K, Niemitalo O, Casteleijn MG, Juffer AH & Neubauer P (2011) High-temperature cultivation and 5′ mRNA optimization are key factors for the efficient overexpression of thermostable *Deinococcus geothermalis* purine nucleoside phosphorylase in *Escherichia coli*. *J Biotechnol* **156**, 268–274.
28 Ukkonen K, Vasala A, Ojamo H & Neubauer P (2011) High-yield production of biologically active recombinant protein in shake flask culture by

combination of enzyme-based glucose delivery and increased oxygen transfer. *Microb Cell Fact* **10**, 107.
29 Cacciapuoti G, Fuccio F, Petraccone L, Del Vecchio P & Porcelli M (2012) Role of disulfide bonds in conformational stability and folding of 5′-deoxy-5′-methylthioadenosine phosphorylase II from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochim Biophys Acta* **1824**, 1136–1143.
30 Jorda J & Yeates TO (2011) Widespread disulfide bonding in proteins from thermophilic Archaea. *Archaea* **2011**, 1–9.
31 Beeby M, O'Connor BD, Ryttersgaard C, Boutz DR, Perry LJ & Yeates TO (2005) The genomics of disulfide bonding and protein stabilization in thermophiles. *PLoS Biol* **3**, 1549–1558.
32 Ferreira AC, Nobre MF, Rainey FA, Silva MT, Wait R, Burghardt J, Chung AP & da Costa MS (1997) *Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *Int J Syst Bacteriol* **47**, 939–947.
33 Suzuki Y, Kishigami T, Inoue K, Mizoguchi Y, Eto N, Takagi M & Abe S (1983) *Bacillus thermoglucosidasius* sp. nov., a new species of obligately thermophilic bacilli. *Syst Appl Microbiol* **4**, 487–495.
34 Sako Y, Nomura N, Uchida A, Ishida Y, Morii H, Koga Y, Hoaki T & Maruyama T (1996) *Aeropyrum pernix* gen nov, sp nov, a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 degrees C. *Int J Syst Bacteriol* **46**, 1070–1077.
35 Li X, Jiang X, Li H & Ren D (2008) Purine nucleoside phosphorylase from *Pseudoalteromonas* sp. Bsi590: molecular cloning, gene expression and characterization of the recombinant protein. *Extremophiles* **12**, 325–333.
36 Araki T, Ikeda I, Matoishi K, Abe R, Oikawa T, Matsuba Y, Ishibashi H, Nagahara K & Fukuiri Y (2005) United States Patent No. US 6858721 B2.
37 Stepchenko VA, Seela F, Esipov RS, Miroshnikov AI, Sokolov YA & Mikhailopulo IA (2012) Enzymatic synthesis of 2′-deoxy-β-D-ribonucleosides of 8–azapurines and 8–aza-7-deazapurines. *Synlett* **23**, 1541–1545.
38 Bennett EM, Li C, Allan PW, Parker WB & Ealick SE (2003) Structural basis for substrate specificity of *Escherichia coli* purine nucleoside phosphorylase. *J Biol Chem* **278**, 47110–47118.
39 Pugmire MJ & Ealick SE (2002) Structural analyses reveal two distinct families of nucleoside phosphorylases. *Biochem J* **361**, 1–25.
40 Mikleušević G, Štefanić Z, Narczyk M, Wielgus-Kutrowska B, Bzowska A & Luić M (2011) Validation of the catalytic mechanism of *Escherichia coli* purine nucleoside phosphorylase by structural and kinetic studies. *Biochimie* **93**, 1610–1622.
41 Prakash TP & Bhat B (2007) 2′-Modified oligonucleotides for antisense therapeutics. *Curr Top Med Chem* **7**, 641–649.
42 Manoharan M (1999) 2′-Carbohydrate modifications in antisense oligonucleotide therapy: importance of conformation, configuration and conjugation. *Biochim Biophys Acta* **1489**, 117–130.
43 Liu P, Sharon A & Chu CK (2008) Fluorinated nucleosides: synthesis and biological implication. *J Fluor Chem* **129**, 743–766.
44 Heidenreich O, Benseler F, Fahrenholz A & Eckstein F (1994) High activity and stability of hammerhead ribozymes containing 2′-modified pyrimidine nucleosides and phosphorothioates. *J Biol Chem* **269**, 2131–2138.
45 Pieken WA, Olsen DB, Benseler F, Aurup H & Eckstein F (1991) Kinetic characterization of ribonuclease-resistant 2′-modified hammerhead ribozymes. *Science* **253**, 314–317.
46 Tennilä T, Azhayeva E, Vepsäläinen J, Laatikainen R, Azhayev A & Mikhailopulo IA (2000) Oligonucleotides containing 9–(2-deoxy-2-fluoro-β-D–arabinofuranosyl)-adenine and -guanine: synthesis, hybridization and antisense properties. *Nucleosides Nucleotides Nucleic Acids* **19**, 1861–1884.
47 Kalota A, Karabon L, Swider CR, Viazovkina E, Elzagheid M, Damha MJ & Gewirtz AM (2006) 2′–Deoxy-2′-fluoro-beta-D-arabinonucleic acid (2′F-ANA) modified oligonucleotides (ON) effect highly efficient, and persistent, gene silencing. *Nucleic Acids Res* **34**, 451–461.
48 Doskocil J & Holy A (1977) Specificity of purine nucleoside phosphorylase from *Escherichia coli*. *Collect Czech Chem Commun* **42**, 370–383.
49 Lehikoinen PK, Sinnott ML & Krenitsky TA (1989) Investigation of alpha-deuterium kinetic isotope effects on the purine nucleoside phosphorylase reaction by the equilibrium–perturbation technique. *Biochem J* **257**, 355–359.
50 Bzowska A, Kulikowska E, Poopeiko NE & Shugar D (1996) Kinetics of phosphorolysis of 3–(β-D–ribofuranosyl)adenine and 3–(β-D–ribofuranosyl) hypoxanthine, non-conventional substrates of purine-nucleoside phosphorylase. *Eur J Biochem* **239**, 229–234.
51 Dessanti P, Zhang Y, Allegrini S, Tozzi MG, Sgarrella F & Ealick SE (2012) Structural basis of the substrate specificity of *Bacillus cereus* adenosine phosphorylase. *Acta Crystallogr D Biol Crystallogr* **68**, 239–248.
52 Erion MD, Stoeckler JD, Guida WC & Walter RL (1997) Purine nucleoside phosphorylase. 2. Catalytic mechanism. *Biochemistry* **36**, 11735–11748.
53 Palmer T (1985) Understanding Enzymes, 2nd edn. Wiley, New York.

54 Plach H, Westhof E, Lüdemann H-D & Mengel R (1977) Solution conformational analysis of 2′-amino-2′-deoxyadenosine, 3′-amino-3′-deoxyadenosine and puromycin by pulsed nuclear magnetic resonance methods. *Eur J Biochem* **80**, 295–304.
55 Meanwell NA (2011) Synopsis of some recent tactical application of bioisosteres in drug design. *J Med Chem* **54**, 2529–2591.
56 Sivets GG, Kalinichenko EN & Mikhailopulo IA (2006) Synthesis of C2′-beta-fluoro-substituted adenine nucleosides via pivaloyl derivatives of adenosine and 3′-deoxyadenosine. *Lett Org Chem* **3**, 402–408.
57 Sivets GG, Kalinichenko EN & Mikhailopulo IA (2007) Synthesis and conformational analysis of 1′- and 3′-substituted 2-deoxy-2-fluoro-D–ribofuranosyl nucleosides. *Helv Chim Acta* **90**, 1818–1836.
58 Panula-Perälä J, Šiurkus J, Vasala A, Wilmanowski R, Casteleijn M & Neubauer P (2008) Enzyme controlled glucose auto-delivery for high cell density cultivations in microplates and shake flasks. *Microb Cell Fact* **7**, 31.
59 Sambrook J & Russell DW (2001) Molecular Cloning: A Laboratory Manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
60 Desjardins P, Hansen JB & Allen M (2001) Microvolume spectrophotometric and fluorometric determination of protein concentration. In Current Protocols in Protein Science (Coligan JE, Dunn BM, Speicher DW & Wingfield PT, eds), pp. 3.10.1–3.10.16. Wiley, New York.
61 Copeland RA (2000) Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis, 2nd edn. Wiley-VCH, Weinheim.
62 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
63 Notredame C, Higgins DG & Heringa J (2000) T–coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.
64 Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V & Notredame C (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D–coffee. *Nucleic Acids Res* **34**, W604–W608.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Relative pH-dependent activity of the thermostalbe PNPs with natural and modified substrates.

Fig. S2. Spectral chromatograms for synthesized $dAdo_{2'F}$ and $dAdo^{2'F}$.

Table S1. Identification of ApMTAP by MALDI-TOF MS quatered measurements.

Table S2. Peptides identification of ApMTAP by MALDI-TOF MS and MS/MS.

Table S3. pH-dependent activity of the thermostalbe PNPs with natural and modified substrates.

Table S4. The structural and electronic features of Ado, $dAdo_{2'NH2}$, $dAdo_{2'F}$ and $dAdo^{2'F}$.

Supporting information

Table S1. Identification of ApMTAP by MALDI-TOF MS quatered measurements.

Table S2. Peptides identification of ApMTAP by MALDI-TOF MS and MS/MS.

Table S3. pH-dependent activity of the thermostalbe PNPs with natural and modified substrates.

Table S4. The structural and electronic features of Ado, $dAdo_{2'NH2}$, $dAdo_{2'F}$ and $dAdo^{2'F}$.

Fig. S1. Relative pH-dependent activity of the thermostalbe PNPs with natural and modified substrates.

Fig. S2. Spectral chromatograms for synthesized $dAdo_{2'F}$ and $dAdo^{2'F}$.

Table S1. Identification of ApMTAP by MALDI-TOF MS quatered measurements*

Measurement	Peptide Count	Protein Score[a]	Seq. Cov. %
1	15	139.00	64.00
2	12	287.00	62.00
3	7	90.00	24.00
4[b]	12	202.00	50.00

[a] Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits, where ion score is a measurement of how well is the observed MS/MS spectrum matches to the stated peptide.

[b] Matched peptides were shown in **Bold Red** (Score: 202, Expect: 3e-17, Sequence Coverage: 50%):

1 MMR**GSHHHHH HGSRKPVHLE AGPGDVAPLV VAVGDPGRAE RLATGLLEDA**
51 **RLVSSAR**GLK VYTGSFNGSE VTIATHGIGG PSAAVVFEEL R**MLGAEVLVR**
101 **LGTSGGLSKD LR**LGDVVVAA GAGCYWGSGG SIQYAGERPM CLPASPDPIL
151 TAGIYRGLSS R**LGDRVVLAP VMSSDAFYAE TPEAAGR**WRS LGMAAVEMEL
201 HTLFSISWIR GFR**SAGVLIV SDLLLPEGFK RITPGELARR** EVEVGRALLE
251 VLTGG

* The protein spots were excised from stained SDS-PAGE gel (Fig.1B, lane 3). The tryptic digest with subsequent spotting on a MALDI-target was carried out automatically with the Ettan Spot Handling Workstation (Amersham Biosciences, Uppsala, Sweden) using tryptic digests mprotocol. The MALDI-TOF and MALDI-TOF-TOF measurements were carried out on the 4800 MALDI TOF-TOF Analyzer (Applied Biosystems, Foster City, CA, USA). This instrument is designed for high throughput measurement, being automatically able to measure the samples, calibrate the spectra and analyze the data using the 4000 Explorer™ Software V3.5.3. The TOF-MS settings were: a mass range from 900 to 3700 Da, a peak density of 20 peaks per 200 Da, maximal 65 peaks per spot and an S/N ratio of 15. The TOF-TOF-MS settings were: a mass range from 60 to Precursor - 20 Da, a peak density of 50 peaks per 200 Da and maximal 65 peaks per precursor. The peak list was created for an S/N ratio of 10. For data base search, the Mascot search engine Version is 2.1.04 (Matrix Science Ltd, London, UK). http://www.matrixscience.com/

Table S2. Peptides identification of ApMTAP by MALDI-TOF MS and MS/MS

Peptide Number	Calculated Mass	Observed Mass	Match Error (Da)	Match Error (ppm)	Sequence Position		Sequence	MS/MS Ion Score*
					Start	End		
1	1003.5604	1003.5537	-0.0067	-7	92	100	MLGAEVLVR	
2	1012.5897	1012.5716	-0.0181	-18	232	240	ITPGELARR	
3	1058.5841	1058.5657	-0.0184	-17	42	51	LATGLLEDAR	
4	1203.6692	1203.6539	-0.0153	-13	101	112	LGTSGGLSKDLR	
5	1285.5795	1285.569	-0.0105	-8	4	14	GSHHHHHHGSR	66.8300018
6	1285.5795	1285.569	-0.0105	-8	4	14	GSHHHHHHGSR	
7	1671.9388	1671.9205	-0.0183	-11	42	57	LATGLLEDARLVSSAR	
8	1914.1058	1914.0894	-0.0164	-9	214	231	SAGVLIVSDLLLPEGFKR	
9	2281.1169	2281.1287	0.0118	5	166	187	VVLAPVMSSDAFYAETPEAAGR	
10	2297.1118	2297.1218	0.01	4	166	187	VVLAPVMSSDAFYAETPEAAGR	
11	2350.2876	2350.2996	0.012	5	15	38	KPVHLEAGPGDVAPLVVAVGDPGR	120.709999
12	2350.2876	2350.2996	0.012	5	15	38	KPVHLEAGPGDVAPLVVAVGDPGR	
13	2706.4685	2706.5132	0.0447	17	15	41	KPVHLEAGPGDVAPLVVAVGDPGRAER	
14	2738.3455	2738.385	0.0395	14	162	187	LGDRVVLAPVMSSDAFYAETPEAAGR	
15	3165.585	3165.6614	0.0764	24	61	91	VYTGSFNGSEVTIATHGIGGPSAAVVFEELR	

* Ion score is -10 × Log (*P*), where *P* is the probability that the observed match is a random event. Ion score greater than 49 means the ions were statistically significant ($P<0.05$).

Table S3. pH-dependent activity of the thermostalbe PNPs with natural and modified substrates.

Enzyme		$dAdo_{2'NH2}$		$dIno_{2'NH2}$			dAdo			Ado			dIno			Ino	
	pH	Activity [U/mg]	SD	Activity [U/mg]	SD	pH	Activity [U/mg]	SD	pH	Activity [U/mg]	SD	pH	Activity [U/mg]	SD	pH	Activity [U/mg]	SD
DgPNP	6.3	0.447	0.01	0.040	0.001	5.6	148.85	9.50	5.9	112.24	14.07	5.6	88.04	4.03	5.6	31.34	4.78
	7.0	1.039	0.11	0.045	0.005	7.0	210.92	14.61	7.0	238.19	18.96	7.0	86.11	5.00	7.0	36.62	6.17
	8.3	0.376	0.01	0.015	0.004	8.5	88.50	0.36	8.2	180.22	0.41	8.5	40.53	0.65	8.5	18.40	0.66
GtPNP	6.3	2.743	0.35	1.972	0.492	5.6	286.41	18.87	5.9	250.19	48.94	5.6	415.58	37.80	5.6	167.25	9.07
	7.0	8.353	0.14	1.958	0.401	7.0	237.13	21.60	7.0	410.29	12.34	7.0	285.38	6.42	7.0	192.34	13.67
	8.3	0.877	0.13	0.153	0.042	8.5	83.71	13.08	8.2	340.78	1.65	8.5	88.81	0.22	8.5	70.55	7.07
ApMTAP	6.3	6.708	0.74	4.359	0.279	5.6	114.57	4.83	5.9	21.72	0.38	5.6	302.37	9.61	5.6	34.07	2.86
	7.0	3.950	0.26	7.615	0.861	7.0	96.25	1.16	7.0	42.46	1.67	7.0	154.93	14.34	7.0	45.42	0.90
	8.3	0.207	0.07	3.318	0.232	8.5	7.59	2.85	8.2	34.58	2.93	8.5	43.51	2.17	8.5	66.95	2.07

Reaction conditions: 1 mM substrate in 50 mM potassium phosphate buffer (pH 5.6~8.5) at 55°C (DgPNP), 70°C (GtPNP), and 80°C (ApMTAP). Reactions were stopped with cold MeOH in the initial reaction phase.

Table S4. The structural and electronic features a of Ado, $dAdo_{2'NH2}$, $dAdo_{2'F}$ and $dAdo^{2'F}$

Compound	Base conformation χ, deg [b]	β-D-Pentofuranose conformation	Selected data					
			Bond length C1′-N9 (Å)	Electronic charge (ε)				Others
				C1′	N9	O4′	N7	
Ado	-114.6; *anti* *(-anticlinal)*	*S*; $_3E$ (C3′-*exo*) *P* = 189.5°, υ_{max} = 26.0°	1.47683	0.166	-0.272	-0.249	-0.267	C2′-OH: -0.098 ε C3′-OH: -0.095 ε
Ado [c]	-171.1; *anti* *(antiperiplanar)*	*N*; ^3E (C3′-*endo*) *P* = 7.2°, υ_{max} = 36.0°	1.466					
$dAdo_{2'NH2}$	-117.6; *anti* *(-anticlinal)*	*E*ast; $^0T_1$ (O4′-*endo*) *P* = 102.5°, υ_{max} = 38.9°	1.46766	0.178	-0.261	-0.260	-0.261	C2′-NH_2: -0.066 ε
$dAdo_{2'NH3^{\oplus}}$ Cl^-	-178.8; *anti* *(antiperiplanar)*	*S*; ^2E (C2′-*endo*) *P* = 165.5°, υ_{max} = 43.0°	1.48539	0.180	-0.268	-0.248	-0.257	C2′-$NH_3^{\oplus}$: +0.402 ε
$dAdo_{2'F}$	-113.4; *anti* *(-anticlinal)*	*S*; $_3E$ (C3′-*exo*) *P* = 203.0°, υ_{max} = 25.7°	1.47357	0.161	-0.268	-0.251	-0.266	C2′-F: -0.139 ε
$dAdo_{2'F}$ [d]	-164.3; *anti* *(antiperiplanar)*	*N*; ^3E (C3′-*endo*) *P* = 8.8°, υ_{max} = 35.7°	1.468					
$dAdo^{2'F}$	-118.6; *anti* *(-anticlinal)*	*S*; $_3E$ (C3′-*exo*) *P* = 206.5°, υ_{max} = 20.3°	1.47249	0.162	-0.267	-0.258	-0.269	C2′-F: -0.137 ε

[a] The structures have been analyzed by the restricted Hartree–Fock (RHF) *ab initio* method using basis set of STO-3G (Hyper-Chem, 8.1 release) and the PM3 geometry optimization as starting approximation for the *ab initio* calculations.

[b] The χ angle is defined as torsion angle O4′-C1′-N9-C4 (*anti:* -180°<χ <-90°; *high-anti:* -90° <χ <-70°*; anticlinal*: -150°<χ <-90°; *antiperiplanar*: -180°<χ <-150°; *syn:* χ = 0° ± 90°; *high-syn*: 90° <χ <110°); an *anti* conformation of adenine around the C1′-N9 glycosidic bond was found for all the nucleosides listed above except for the $NH_3^{\oplus}$ form of $dAdo_{2'NH2}$, for which an *antiperiplanar* conformation resulted from calculations.

[c] X-ray data taken from Lai, T.F.; Marsh, R.E (1972) *Acta Cryst. Sect. B*, **B28**, 1982-1989; see also: Klooster, W.T.; Ruble, J.R.; Craven, B.M.; McMullan, R.K (1991) *Acta Cryst. Sect. B*, **B47**, 376-383.

[d] X-ray data taken from Morishita, K.; Hakoshima, T.; Fujiwara, T.; Tomita, K-i.; Kaneyasu, T.; Uesugi, S.; Ikehara, M (1984) *Acta Cryst. Sect. C*, **C40**, 434-436.

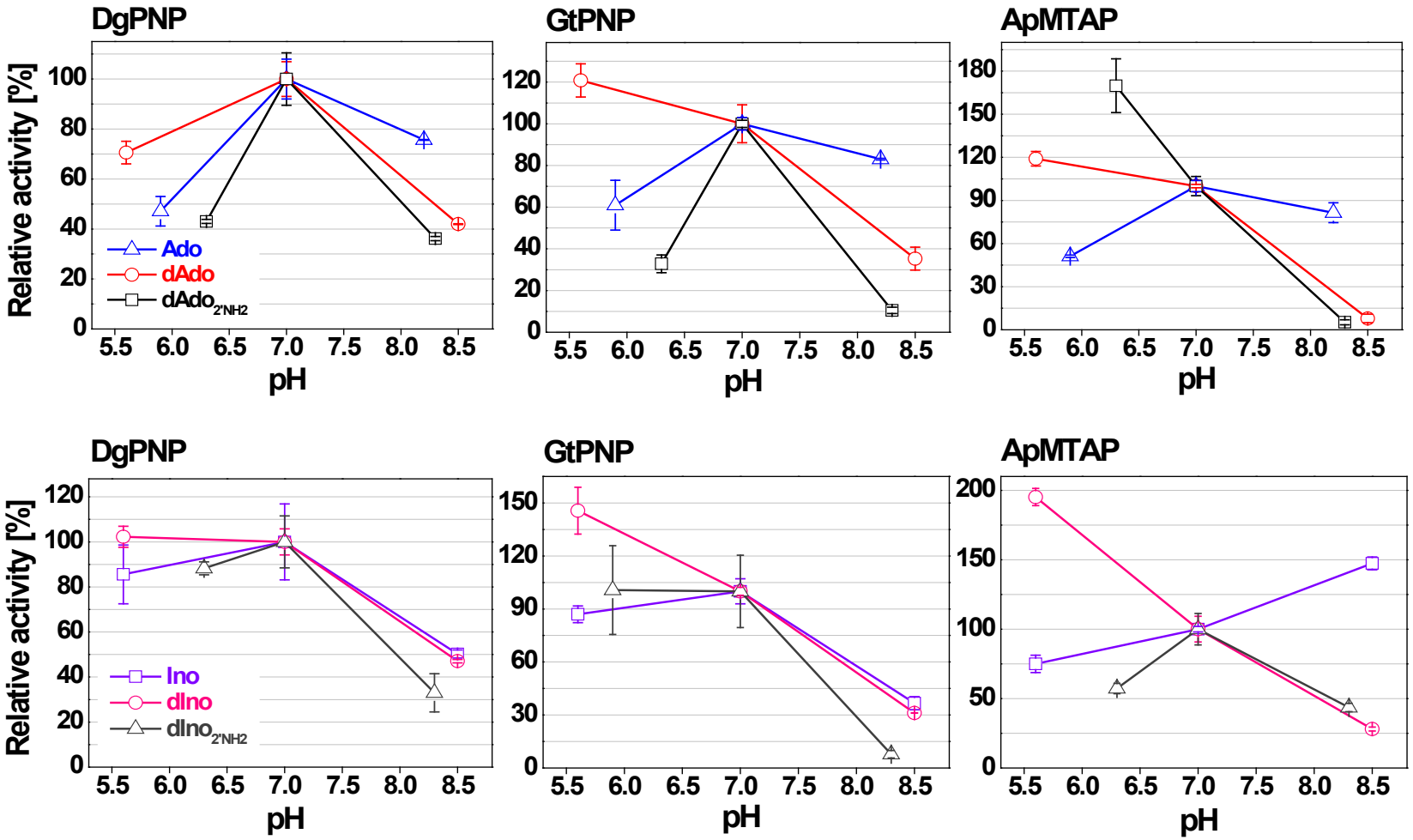
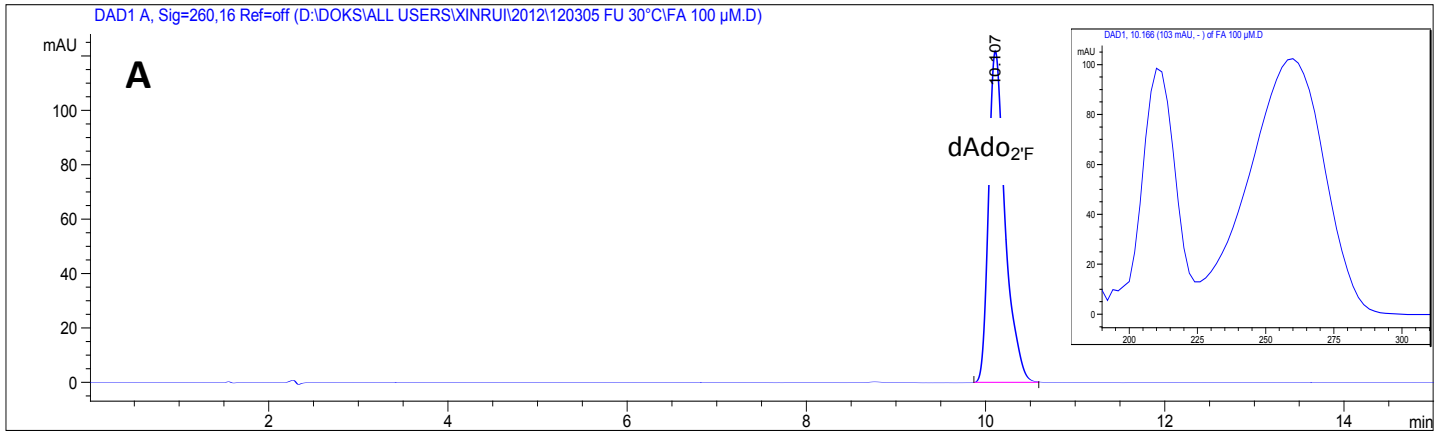
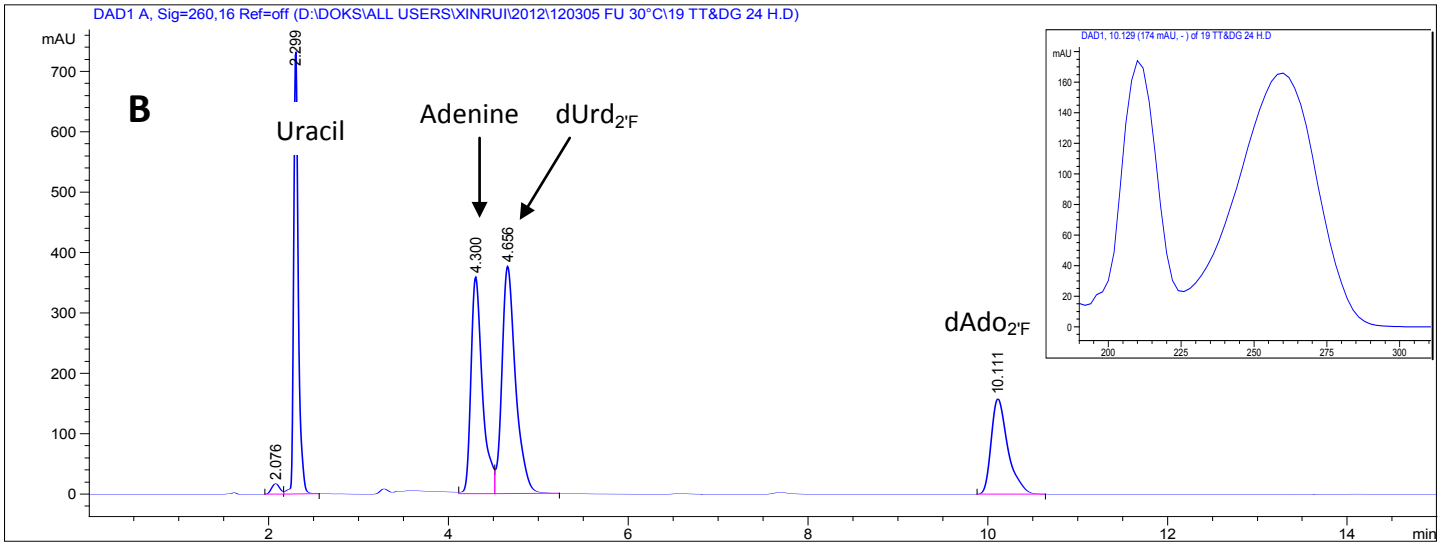


Fig. S1. Relative pH-dependent activity of the thermostalbe PNPs with natural and modified substrates (the activity of every PNP with each substrate at pH 7 was set to 100 %). Reaction conditions (see Table S3)


DAD1 A, Sig=260,16 Ref=off (D:\DOKS\ALL USERS\XINRUI\2012\120305 FU 30°C\FA 100 µM.D)
A
mAU
10.107
dAdo2'F
min


DAD1 A, Sig=260,16 Ref=off (D:\DOKS\ALL USERS\XINRUI\2012\120305 FU 30°C\19 TT&DG 24 H.D)
B
mAU
2.299
Uracil
Adenine
dUrd2'F
4.300
4.656
dAdo2'F
10.111
2.076
min

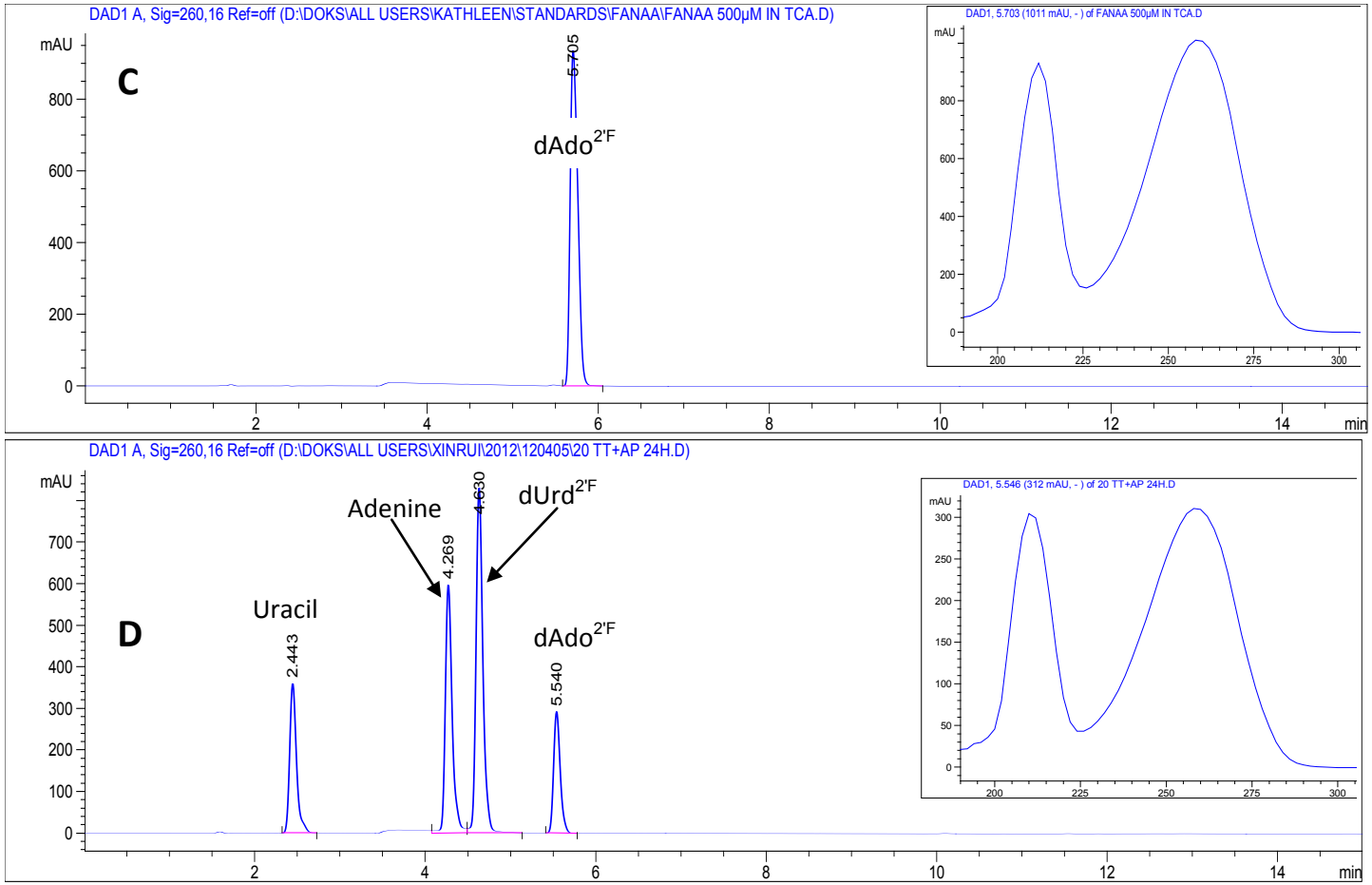


Fig. S2. Spectral chromatograms for synthesized $dAdo_{2'F}$ and $dAdo^{2'F}$. HPLC chromatogram of 100 µM $dAdo_{2'F}$ (A), $dAdo_{2'F}$ synthesis mixture (B), 500 µM $dAdo^{2'F}$ (C), and $dAdo^{2'F}$ synthesis mixture (D). Inserted UV spectrum at 10.2 min ($dAdo_{2'F}$) (A), 10.1 min ($dAdo_{2'F}$) (B), 5.7 min ($dAdo^{2'F}$) (C), and 5.5 min ($dAdo^{2'F}$) (D).

Paper III

(see Chapter 4)

Paper IV

Enzymatic Synthesis of 2,6-Dihalogenated Purine Nucleosides

Xinrui Zhou,[a] Kathleen Szeker,[a] Linyu Jiao,[b] Igor A. Mikhailopulo,[c,*] and Peter Neubauer[a,*]

[a] Chair of Bioprocess Engineering, Institute of Biotechnology, Technische Universität Berlin, Ackerstraße 76 ACK24, D-13355 Berlin, Germany
Phone: (+49)-030-3147-2269; fax: (+49)-030-3142-7577; email: peter.neubauer@tu-berlin.de

[b] Institute of Chemistry, Technische Universität Berlin, Straße des 17. Juni 115, D-10623, Germany

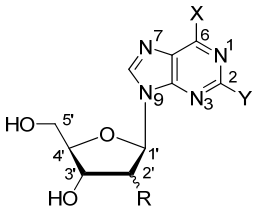
[c] Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Acad. Kuprevicha 5/2, 220141 Minsk, Belarus
Phone: (+375)-17267-8148; fax: (+375)-17267-8161; email: imikhailopulo@gmail.com

Introduction

Nucleoside analogues are broadly used as anti-tumour and anti-viral agents. In the living cell, they manifest biological effects via diverse metabolic transformations, primarily as a result of conversion into derivatives of phosphoric acid, or themselves by inhibiting the metabolic pathways or as blockers of the synthesis of nucleic acids.[1]

Table 1. Structure and bioactivity of halogenated purine nucleosides.

	Nucleoside	X	Y	R	Bioactivity
1	Cladribine	NH_2	Cl	H	anti leukemia; used in clinic[2]
2	2Cl-Ado	NH_2	Cl	OH (*ribo*)	induce apoptosis in several cells[3]
3	Clofarabine	NH_2	Cl	F (*arabino*)	anti leukemia; used in clinic[4]
4	2F-dAdo	NH_2	F	H	cytotoxic[5]
5	2F-Ado	NH_2	F	OH (*ribo*)	highly cytotoxic[5-6]
6	Fludarabine	NH_2	F	OH (*arabino*)	anti leukemia; used in clinic[7]
7	26DCP-R	Cl	Cl	OH (*ribo*)	smooth muscle relaxant[8]
8	26DCP-dR	Cl	Cl	H	unknown
9	6C2FP-R	Cl	F	OH (*ribo*)	unknown
10	6C2FP-dR	Cl	F	H	unknown



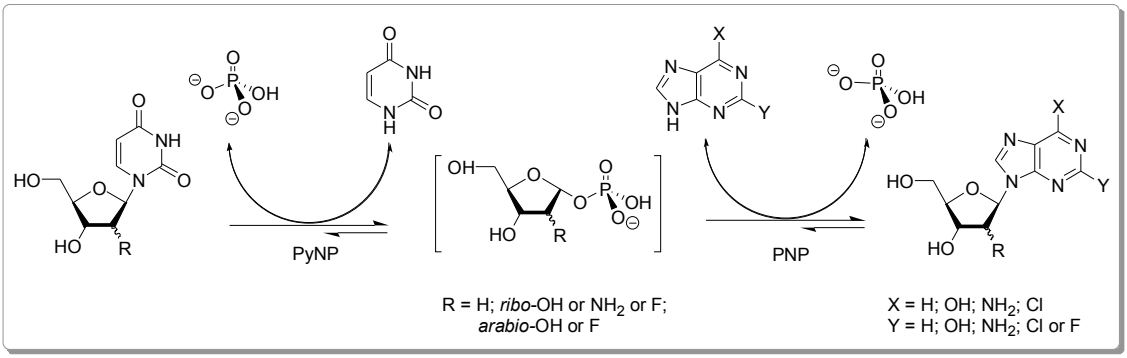
Among the great diversity of the biologically important nucleosides of this family, adenosine analogues containing chlorine (**1-3**) or fluorine (**4-6**) atom at the C2 carbon atom as well as 2,6-dichloro- and 6-chloro-2-flouropurine ribo- (**7** and **9**) and 2′-deoxyribo-nucleosides (**8** and **10**) are of particular interest. Indeed, (*i*) Cladribine (**1**), Clofarabine (**3**) and Fludarabine (**6**) attract much attention during last two decades as drugs of great therapeutic potential, in particular for the treatment of leukemia[2, 4, 7]; (*ii*) 2-chloroadenosine (2Cl-Ado, **2**) was shown to be an agonist of the adenosine receptor that induces apoptosis in several cell lines[3]; (*iii*) 2′-deoxy-2-fluoroadenosine (2F-dAdo, **4**) and 2-fluoroadenosine (2F-Ado, **5**) are the first members of this group of adenosine analogues, have been synthesized in the middle of the last century, they showed high cytostatic activity against tumour cells that, unfortunately, is accompanied by the high toxicity for normal cells[5-6]. Di-halogenated purine nucleosides **7**-**10** (see Table 1

Table) are of great value first of all as precursors for the synthesis of numerous purine nucleosides in the framework of the R&D of efficient routes to the preparation of known substances as well as new nucleosides of potential biological importance.

Until now, the vast majority of the considered nucleoside analogues have been synthesized by chemical methods. Two approaches to the synthesis of purine nucleosides have been widely and extensively studied. The first, convergent approach consists in the condensation of purine base with a suitable pentofuranose derivative and subsequent deprotection (see, *e.g.* [5-6, 9]; and the reviews [10]). The second, comprises in the chemical transformation of purine base or/and pentofuranose fragment of the natural commercially available *ribo*- or 2′-*deoxyribo*- nucleoside

in the desired compound [11]. Despite the very impressive progress achieved in the development of chemical methods, the preparation of purine nucleosides remains a challenging problem. Without going into details, it should be noted that a common drawback of both approaches is indispensable introduction and subsequent removal of the protective groups, which is associated with chromatographic purification on almost every step, a lack of selectivity of chemical reactions, and the problems associated with the poor *regio*- and *stereo*- selectivity of chemical reactions (see, *e.g.* [12] and the papers cited therein).

In complete contrast, enzymatic synthesis of nucleosides by the transglycosylation reaction proceeds strictly *stereo*- and *regio*- selectively (except for a few specific substrates[13]), which requires organic solvents (ethanol, acetonitrile) only on steps of isolating of the individual desired compound. Obviously, the substrate specificity of the enzyme and physicochemical properties of the substrates are critically important factors of this reaction. From a viewpoint of the practical synthesis, nucleoside phosphorylases (NPs) of diverse origin are presently very useful biocatalysts for the synthesis of base and pentofuranose modified nucleosides[13-14], which include pyrimidine NPs (PyNP, EC 2.4.2.2; UP, EC 2.4.2.3; TP, EC 2.4.2.4) and purine NPs (PNP, EC 2.4.2.1; MTAP, EC 2.4.2.28). These enzymes catalyze the reversible phosphorolysis of nucleoside in the presence of inorganic phosphate to afford α-D-ribofuranose-1-phosphate (α-D-Rib-1P) and nucleobase. The basic synthetic strategy is illustrated in Scheme 1.



Scheme 1. Enzymatic synthesis of modified nucleosides.

The rationality of the transglycosylation reaction consists in the transfer of the pentofuranose moiety of a donor, which is commercially available pyrimidine nucleoside or its pentofuranose modified derivative, to the purine base; the direction of the reaction is determined by the differences in the equilibrium of nucleoside phosphorolysis catalyzed by purine *vs* pyrimidine phosphorylases. The equilibrium of the phosphorolysis is usually shifted to the nucleoside formation, which is more pronounced for the reactions catalyzed by PNP than that by PyNP, thus the coupled reactions proceed in the direction of purine nucleoside synthesis. Moreover, it is effective to use pyrimidine nucleosides as pentofuranose donor because the

product (pyrimidine base) from the first step will not become the competitor of the reactant (purine base) for the second step (Scheme 1).

A number of halogenated nucleosides have been synthesized by NPs from different sources. For instance, 2-Cl (or F) substituted adenine is accepted as substrate of PNPs from *E. coli* [15], *Geobacillus stearothermophilus*[16] and *Sulfolobus solfataricus*[17] for the synthesis of Cladribine **1** [15a-d, 16], Fladarabine **6** [15e, 16] and their analogues; 6-Cl substituted purine is accepted by PNPs from *E. coli* [18] and *Aeromonas hydrophila*.[19] The reported conversions resulted in satisfactory yields of desired nucleosides (mainly HPLC data), and some of them gave yields in the range from 50 % to 80 % for the pure individual nucleosides. However, the wider application was restricted due to the following limitations:

Firstly, broadly used bacterial whole cells as biocatalysts possess a number of enzymes that can consume substrates, catalyze undesired transformation of substrates or formation of nucleosides and, moreover, secrete in a medium of cell metabolite(s). The purified NPs (*e.g., E. coli* enzymes) typically exhibit lower activity towards the analogues of the natural substrates, and they will become inactive under the harsh conditions including high temperature, high substrate concentration and organic solvents. Even though today molecular evolution techniques are available to improve the *E. coli* enzymes[20] or even adapting the cell to organic solvents[21] , the genetic reservoir of the variety of microorganisms probably offers better starting points. Enzymes from thermophiles are advantageous over the conventional bacterial enzymes for the synthesis of modified nucleosides. Abstracting from their ability to operate at higher temperatures, *a priori* can be expected a wider range of substrate specificity due to their evolutionary distance to *E. coli*.

Secondly, the purine bases 2-chloroadenine (2Cl-Ade) and 2-fluoroadenine (2F-Ade) have extremely low solubility in aqueous solutions which are most suitable for the enzymatic reactions. We tested the effect of temperatures on the solubility of 2Cl-Ade and 2F-Ade as well as relevant compounds in water solution in the range from 25 °C to 80 °C and found that the use of halides 26DCP and 6C2FP as substrates of the enzymatic synthesis is much more promising than 2Cl-Ade and 2F-Ade from the viewpoint of preparative synthesis because of their greater solubility (Table 2). Nevertheless, the use of 2Cl-Ade and 2F-Ade in the enzymatic synthesis of nucleosides was described in a number of publications (*e.g.* [15b, 16]).

Table 2. Temperature dependence of the substrate solubility in aqueous solution.[a]

Temp. [°C]	2CA	2FA	Ade	26DCP [mM][b]	6C2FP	6C2FP-R
25	0.1	0.9	5.7[c]	15.8[c]	29.0[c]	79.9
40	0.2	1.4	9.9[d]	17.6	49.3	211.5
55	0.3	2.1	16.5	19.8	53.4	ND[e]
65	0.4	2.2	21.7	34.0	81.8	ND
80	0.5	4.0	36.9	38.4	150.2	ND

[a] Solubility was measured by UV spectroscopy at the maximum wavelength of each compounds (284-288 nm) in 2 mM potassium phosphate buffer (pH 7.0). 2CA= 2Cl-Ade; 2FA= 2F-Ade; Ade= adenine; 26DCP= 2,6-dichloropurine; 6C2FP= 6-chloro-2-fluoropuine; 6C2FP-R= 6-chloro-2-fluoropuine riboside.

[b] The largest relative standard deviation is less than 3 % (n = 4).

[c] Determined at 22 °C.

[d] Determined at 43 °C.

[e] ND= not determined.

Results and Discussion

Scrutiny of the published methods for the chemo-enzymatic synthesis of base and pentofuranose modified purine nucleosides prompted us to formulate a new versatile approach to the synthesis of this class of biologically important compounds. This approach consists in the enzymatic synthesis of 2,6-dihalogenated purine nucleosides, in particular 2,6-dichloropurine (26DCP) and 6-chloro-2-fluoropurine (6C2FP), as valuable scaffolds for the enzymatic and/ or chemical preparation of a broad palette of modified nucleosides. The motivation of this work is associate with the preparation of a number of PyNPs and PNPs from the thermophilic microorganisms (TtPyNP from *Thermus thermophilus*, GtPyNP and GtPNP from *Geobacillus thermoglucosidasius*, DgPNP from *Deinococcus geothermalis*, and ApMTAP from *Aeropyrum pernix*) and the study of their substrate properties,[22] which allow us to choose the most efficient biocatalysts for the synthesis of halogenated purine nucleosides. Note that dihalogenated purines have not found application as substrates of the enzymatic synthesis of their nucleosides owing to very poor substrate activity for the other reported PNPs.

The present study was initiated by the comparative testing of activity of DgPNP, GtPNP and ApMTAP under optimal reaction conditions described earlier.[22] With this aim in view, two-step analysis was used, i.e. phosphorolysis of uridine by TtPyNP or GtPyNP to generate α-D-Rib-1P in step 1. After this a purine base and PNP were added to the reaction mixture (step 2) and the formation of the corresponding purine nucleoside was monitored by HPLC. The activity of 2,6-dihalogenated purines was compared with 2,6-diaminopurine (DAP) that is known to be one of the best purine substrate for PNPs from diverse sources. The results are shown in Table 3. As expected, DAP showed best substrate properties for the all PNPs tested; GtPNP showed

superior activities *vs* those of DgPNP and ApMTAP towards the tested purine bases and it was selected for the further study.

Table 3. Apparent activity of PNPs towards modified purine base.

Purine base[b]	Apparent activity[a] [U mg^{-1}]		
	DgPNP 55 °C	GtPNP 70 °C	ApMTAP 80 °C
DAP	98.96	152.76	10.62
6C2FP	0.33	9.81	0.25
26DCP	0.91	4.30	0.33

[a] The apparent activity was determined under the conditions that the ratio of α-D-Rib-1P to purine base was 0.6 mM: 1 mM and the base conversion rate was controlled in the linear range. 1 U = the amount of the enzyme converting 1 μmol of purine base per minute under the defined conditions.

[b] In 2 mM phosphate buffer (pH 7.0). DAP = 2,6-diaminopurine; 6C2FP = 6-chloro-2-fluoropuine; 26DCP = 2,6-dichloropurine.

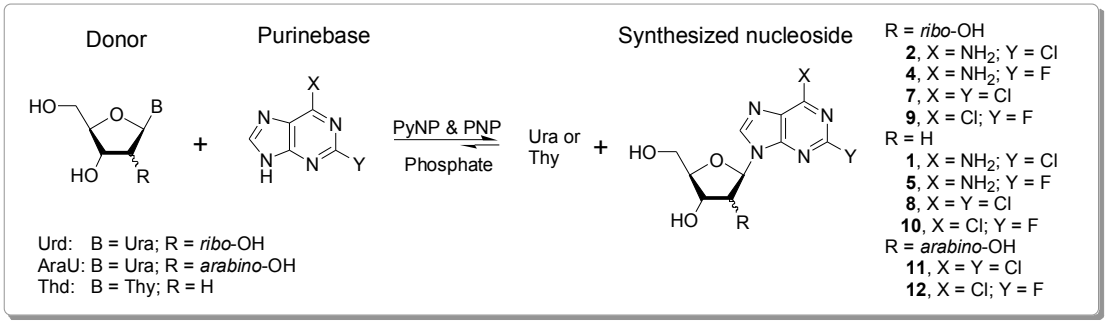
In the next series of experiments, the enzymatic syntheses of ribosides (entries 1-5), 2'-deoxyribosides (entries 6-10), and arabinosides (entries 11-13) using GtPNP and TtPyNP or GtPyNP as biocatalysts were investigated, and the results are summarized in Table 4. The choice of combinations of the enzymes as well as the reaction temperatures are based on the earlier published data of the substrate specificity of the used enzymes.[22] At indicated temperatures, TtPyNP and GtPyNP are interchangeable for the phosphorolysis of uridine and thymidine; however, AraU showed no substrate activity for GtPyNP, but is a substrate for TtPyNP (Table 4).

Analysing the data of Table 4, the excellent substrate properties of the bases studied in the synthesis of *ribo*- and *2'-deoxyribo* nucleosides should be underlined first of all. Indeed, the experimental conditions were not optimized in regard to the donor: acceptor ratio and the reaction time (especially in the case of halogenated bases) and none the less the conversion of 26DCP and 6C2FP reached 54-69 % demonstrating a great potential of the suggested approach to the synthesis of the corresponding nucleosides on the preparative scale. The excellent substrate activity of DAP towards PNP's of diverse origin is well documented [1b, 13-14], however, a high rate of the 2Cl-Ade and 2F-Ade conversion in the corresponding *ribo*- and *2'-deoxyribo* nucleosides is somewhat unexpected taking into account extremely poor solubility of the bases in aqueous media. These data on the whole allow suggesting unusually high efficiency of the GtPNP catalyzed practically irreversible coupling of a base with the intermediary formed α-D-Rib-1P and α-D-dRib-1P promoting the solution of 2Cl-Ade and 2F-Ade in the reaction mixture.

In the case of the synthesis of *ara*-nucleosides, the low phosphorolytic efficiency of TtPyNP along with lower catalytic power of GtPNP toward α-D-ara-1P *vs* that in regard of the natural α-D-Rib-1P and α-D-dRib-1P are obviously the main reason of a low rate of the reaction. A

careful optimization of the reaction conditions is necessary to improve the formation rate and yield of the desired nucleosides.

Table 4. Synthesis of modified purine ribonucleosides *via* enzymatic transglycosylation.[a]



Entry	Pentofuranosyl		Purine Base[c]		Synthesized nucleoside	Time/Temp [h/°C]	Conversion [%]	
	Donor[b]	R	X	Y			Donor[d]	Base[e]
1	Urd	OH (*ribo*)	NH_2	NH_2	DAP-R	1/ 70	66.9	95.3
2			NH_2	Cl	2Cl-Ado (**2**)	1/ 70	45.8	93.3
3			NH_2	F	2F-Ado (**4**)	1/ 70	73.3	91.6
4			Cl	Cl	26DCP-R (**7**)	0.5/ 65	52.8	53.9
5			Cl	F	6C2FP-R (**9**)	0.5/ 65	54.1	63.0
6	Thd	H	NH_2	NH_2	DAP-dR	1/ 70	66.8	80.0
7			NH_2	Cl	Cladribine (**1**)	1/ 70	55.0	80.0
8			NH_2	F	2F-dAdo (**5**)	1/ 70	83.6	65.9
9			Cl	Cl	26DCP-dR (**8**)	0.5/ 65	58.6	67.4
10			Cl	F	6C2FP-dR (**10**)	0.5/ 65	59.7	69.3
11	AraU	OH (*arabino*)	NH_2	NH_2	DAP-araR	3/ 70	54.6	39.5
12			Cl	Cl	26DCP-araR (**11**)	3/ 70	40.8	9.9
13			Cl	F	6C2FP-araR (**12**)	3/ 70	40.0	7.2

[a] Enzymes: TtPyNP 0.1 mg mL^{-1} (5.3 U mL^{-1} towards Urd at 70 °C) and GtPNP 0.2 mg mL^{-1} (2.0 U mL^{-1} towards 6C2FP at 70 °C) (entries 1-3, 6-8, 11-13); GtPyNP (5.1 U mL^{-1} towards Urd at 65 °C) and GtPNP (0.9 U mL^{-1} towards 6C2FP at 65 °C), 0.1 mg mL^{-1} each (entries 4-5, 9-10); Medium: 2 mM *K*-phosphate buffer (pH 7.0).

[b] Pentofuranosyl donor: 2 mM. Urd = uridine, Thd =thymidine, AraU =1-(β-D-arabinofuranosyl)uracil.

[c] Purine base: 1 mM. 2Cl-Ade (entries 2 and 7) can only reach 0.6 mM.

[d] Yield (according to HPLC) based on pyrimidine nucleosides (donor) conversion.

[e] Yield (according to HPLC) based on purine base conversion.

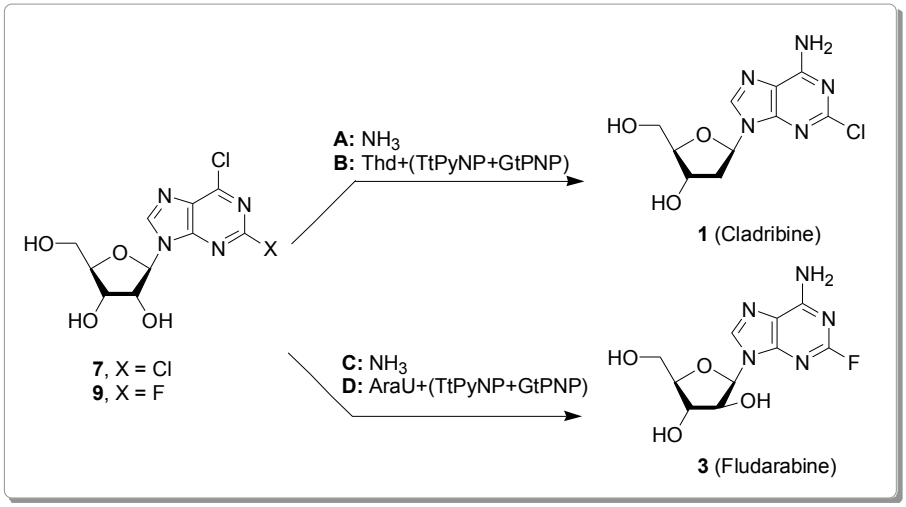
With this aim in view, TtPyNP and GtPNP were immobilized on MagReSyn™ epoxide beads (polyethyleneimine-based iron oxide magnetic microspheres with epoxide functional groups; were kindly gifted from ReSyn Biosciences, Pretoria, South Africa) as one of possible routes to improve the yield of the desired nucleosides. The activity of the immobilized enzymes was measured according to the activity assay as previously described [22] and the following data were obtained: TtPyNP – 4.8 mg enzyme was loaded per mL beads, 41 % residual activity;

GtPNP – 16.5 mg enzyme was loaded per mL beads, 83 % residual activity. It was found that Urd and 6C2FP in concentrations up to 80 mM and 55 mM, respectively, does not inhibit the immobilized enzymes.

In the small scale (0.4 mL), under optimal conditions, the conversion of 26DCP and 6C2FP in the corresponding *ribo*-nucleosides **7** and **9** were found by HPLC to be 78% and 85%, respectively. See "Experimental Section" for more details.

To validate the efficiency of the immobilized enzymes for the preparation of pure nucleosides on the multi milligram scale, the synthesis of 6-chloro-2-fluoro-9-(β-D-ribofuranosyl)-purine (**9**) was studied. The reaction mixture (in 2 mM K-phosphate buffer, pH 7.0; total volume 50 mL) containing 50 mM Urd, 25 mM 6C2FP was kept at 60 °C for **20** h monitoring by HPLC. During this period, the formation of the nucleoside **9** reached 75% and the reaction mixture was worked-up (removing of the biocatalyst by magnet, silica gel column chromatography) to give the desired product in 60% yield of 98 % purity (HPLC).

The results of the synthesis of the nucleoside **9** prompted us to extend this work to the preparation of nucleosides **7**, **8** and **10** aiming at the R&D of the new efficient chemo-enzymatic routes to known and new modified purine nucleosides. This work is in progress. Emphasis is made on the role of the successional treatments for the efficient synthesis of purine nucleosides, *e.g*., Cladribine (**1**) and Fluarabine (**3**), using either ribonucleosides **7** and **9** (first B then A, and first D then C) or their corresponding 2-chloro- and 2-fluoro-adenine derivatives (first A then B, and first C then D) as donors of the relevant purine bases in the transglcosylation reactions (Scheme 2. Chemo-enzymatic routes for the synthesis of Cladribine and Fludarabine.Scheme 2) (*cf*. refs [23]).



Scheme 2. Chemo-enzymatic routes for the synthesis of Cladribine and Fludarabine.

In summary, possibility of application of the tandem of GtPNP & TtPyNP, GtPNP & GtPyNP, and GtPNP and TtPyNP as biocatalysts for the synthesis of corresponding *ribo*, *2'*-

deoxyribo and *arabino* purine nucleosides was studied and it was found that DAP, 2Cl-Ade, 2F-Ade, 26DC-P and 6C2FP are effectively transformed into the corresponding *ribo* and *2'-deoxyribo* nucleosides; the formation of the *arabino* nucleosides of DAP, 26DC-P and 6C2FP proceeded with lesser efficiency and the careful optimization of the reaction conditions is necessary to develop the practical method of their preparation. Immobilisation of GtPNP and TtPyNP on MagReSyn^{TM} Epoxide beads resulted in more efficient transformation of 26DC-P and 6C2FP into the corresponding *ribo* nucleosides as it was exemplified by the synthesis of 6-chloro-2-fluoro-9-(β-D-ribofuranosyl)-purine (**9**) on a multi milligram scale. The results of the present study *in toto* represent a great interest for the R&D of efficient methods of base and pentofuranose modified purine nucleosides.

Experimental Section

Solubility Test

Calibration curves were made by UV spectroscopy for each compound to be tested. The wavelengths of maximum absorption are: 285 nm (2Cl-Ade, $\varepsilon = 624.6\ M^{-1}\,cm^{-1}$), 284 nm (2F-Ade, $\varepsilon = 284.9\ M^{-1}\,cm^{-1}$), 284 nm (Ade, $\varepsilon = 261.5\ M^{-1}\,cm^{-1}$), 288 nm (26DCP, $\varepsilon = 3215\ M^{-1}\,cm^{-1}$), 285 nm (6C2FP, $\varepsilon = 2335\ M^{-1}\,cm^{-1}$; 6C2FP-R, $\varepsilon = 1627\ M^{-1}\,cm^{-1}$). The saturated solutions suspended in 2 mM phosphate buffer (pH 7.0) were incubated at different temperatures in a water bath or on a thermomixer for 30 min. The supernatant was withdrawn and cooled down shortly, then diluted to the suitable concentration for the absorption measurement. Solubility was calculated according to the saturation concentration at the defined temperature.

Expression and Preparation of Enzymes

TtPyNP, GtPyNP, GtPNP, DgPNP and ApMTAP were expressed in *E. coli* BL21 and purified by heat-treatment and standard Ni-NTA affinity chromatography as previously described.[22] Purified TtPyNP has a specific activity of 53 U mg^{-1} at 70 °C and GtPyNP 51 U mg^{-1} at 65 °C towards Urd in 50 mM phosphate buffer (pH 7).

Procedure of Determination of PNP Apparent Activity towards Purine Base

The procedure contains two steps: i) 0.5 mL 4 mM Urd was incubated with 0.1 mg GtPyNP or TtPyNP in 2 mM phosphate buffer (pH 7) for 5 min at the temperature of the second step. In the end of the reaction, equilibrium was reached and 30 % Urd was converted to Ura and α-D-Rib-1P. ii) 0.5 mL 2 mM purine base in 2 mM phosphate buffer (pH 7) with PNP (the concentration was selected in a way that the reaction rate was in the linear range) was added into the first step

reaction for 2 min at 55 °C (DgPNP), 65 °C (GtPNP) or 80 °C (ApMTAP). The reaction mixtures were analysed by HPLC (see supporting information). One unit PNP apparent activity was defined as the amount of PNP which catalyses 1 μmol purine base per minute under the aforementioned conditions.

Enzymatic Synthesis of Halogenated Purine Nucleosides

0.40 mL reaction mixture consisted of 2 mM pentofuranosyl donor (Urd, Thd or AraU), 1 mM purine base, 2 mM phosphate buffer (pH 7.0), 0.1 mg mL^{-1} TtPyNP (or GtPyNP) and 0.1–0.2 mg mL^{-1} GtPNP. Reactions were performed on a thermomixer at defined temperatures for 0.5 to 3 h (Table 4) and stopped by adding ice cold water by a 1:1 dilution. Controls (Table 4, entries 2-3, 7-8, 11-13) without enzyme were performed in parallel and showed no product formation. The conversions were calculated from HPLC analysis, which were the percentage of converted donor or base.

Immobilization of TtPyNP and GtPNP

Immobilization and the optimal reaction conditions for the synthesis of 2,6-dichloro-9-(β-D-ribofuranosyl)-purine (**7**) and 6-chloro-2-fluoro-9-(β-D-ribofuranosyl)-purine (**9**). TtPyNP and GtPNP were immobilized on MagReSyn™ epoxide beads (ReSyn Biosciences, Pretoria, South Africa) according to the accompanied manual with a slight modification. Briefly, 0.04 mL suspension beads (1 mg dry beads) were loaded with TtPyNP (0.19 mg; 4.1 U towards Urd at 70 °C) or GtPNP (0.66 mg; 2.4 U and 5.4 U towards 26DCP and 6C2FP at 70 °C, respectively) in a total volume of 0.40 mL binding buffer (50 mM potassium phosphate buffer, pH 6.0) and incubated on the thermomixer shaking at 950 rpm, 50 °C for 4 h. The quenching step was skipped. Beads were washed and stored at 4 °C in 50 mM potassium phosphate buffer (pH 7.0) for further use.

The optimal conditions for the synthesis of the nucleoside **7** were: 0.4 mL reaction containing 20 mM Urd, 7 mM 26DCP, 1.4 mM K-phosphate (pH 7.0), 4.1 U TtPyNP (towards Urd at 70 °C) and 2.4 U GtPNP (towards 26DCP at 70 °C). The reaction mixture was shaken at 1000 rpm for 30 min at 70 °C and HPLC analysis showed a 78% base conversion in the nucleoside **7**.

The optimal conditions for the synthesis of nucleoside **9** were: 0.4 mL reaction containing 50 mM Urd, 20 mM 6C2FP, 2 mM K-phosphate (pH 7.0), 4.1 U TtPyNP (towards Urd at 70 °C) and 5.4 U GtPNP (towards 6C2FP at 70 °C). The reaction mixture was shaken at 1000 rpm for 180 min at 70 °C and HPLC analysis showed an 85% base conversion in the nucleoside **9**.

Synthesis and purification of 6-chloro-2-fluoro-9-(β-D-ribofuranosyl)-purine (9)

Reaction mixture (50 mL) containing uridine (0.61 g, 2.5 mmol), 6-chloro-2-fluoropurine (0.216 g, 1.25 mmol), potassium phosphate (0.1 mmol, pH 7.0) and immobilized biocatalysts (TtPyNP: 8 mg, *ca.* 67 U, towards Urd at 60 °C; GtPNP: 15 mg, *ca.* 87 U, towards 6C2FP at 60 °C) was enclosed in a 50 mL-falcon tube and rotated by an end-over-end rotator at 55-60 °C in a water bath for 20 h. 75 % conversion on the base was determined by HPLC. The biocatalyst was removed by magnet and the product was isolated by the silica gel column chromatography to afford the desired product **9** as a white solid (0.230 g, 0.75 mmol; HPLC purity 98 %) in 60% yield (see the Supporting Information for more details).

Acknowledgements

We are grateful to Dr. Justin Jordaan (ReSyn Biosciences, Pretoria, South Africa) for the kind donation of the MagReSyn™ Epoxide beads as well as useful suggestions for the immobilization experiments. This work is part of the Cluster of Excellence "Unifying Concepts in Catalysis" coordinated by the Technische Universität Berlin. Financial support by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the German Initiative for Excellence is gratefully acknowledged (EXC 314). Igor A. M. is deeply thankful to the A. von Humboldt-Stiftung (Bonn – Bad-Godesberg, Germany) and the Byelorussian Republican Foundation for Fundamental Research (www.fond.bas-net.by; project #X13MC-027) for partial financial support.

References

[1] a) P. Herdewijin, *Modified Nucleosides: in Biochemistry, Biotechnology and Medicine*, Wiley-VCH, Weinheim, **2008**; b) I. A. Mikhailopulo, A. I. Miroshnikov, *Acta Naturae* **2010**, *2*, 36-59; c) W. Sneader, *Drug Discovery: A History*, John Wiley & Sons, London, **2005**.

[2] T. Robak, P. Robak, *Curr. Pharm. Des.* **2012**, *18*, 3373-3388.

[3] a) I. Bellezza, A. Tucci, A. Minelli, *Anticancer Agents Med Chem* **2008**, *8*, 783-789; b) L. Bastin-Coyette, C. Smal, S. Cardoen, P. Saussoy, E. Van Den Neste, F. Bontemps, *Biochem. Pharmacol.* **2008**, *75*, 1451-1460.

[4] P. L. Bonate, L. Arthaud, W. R. Cantrell, Jr., K. Stephenson, J. A. Secrist, III, S. Weitman, *Nat. Rev. Drug Discovery* **2006**, *5*, 855-863.

[5] J. A. Montgomery, K. Hewson, *J. Med. Chem.* **1969**, *12*, 498-504.

[6] J. A. Montgomery, K. Hewson, *J. Am. Chem. Soc.* **1957**, *79*, 4559.

[7] H. Wallen, J. A. Thompson, J. Z. Reilly, R. M. Rodmyre, J. Cao, C. Yee, *PLoS One* **2009**, *4*, e4749.

[8] S. G. McKenzie, R. Frew, H. P. Bar, *Eur. J. Pharmacol.* **1977**, *41*, 183-192.

[9] G. M. Blackburn, M. J. Gait, D. Loakes, D. M. Williams, *Nucleic Acids in Chemistry and Biology*, 3rd ed., Royal Society of Chemistry, Cambridge, **2006**.

[10] a) H. Vorbrüggen, C. Ruh-Pohlenz, in *Organic Reactions*, John Wiley & Sons, Inc., **2004**; b) J. K. Watts, M. J. Damha, *Can J Chem* **2008**, *86*, 641-656.

[11] a) K. A. Watanabe, S. E. Patterson, in *Frontiers in Nucleosides and Nucleic Acids* (Eds.: R. F. Schinazi, D. C. Liotta), IHL Press, **2005**, pp. 3-56; b) K. W. Pankiewicz, *Carbohydr. Res.* **2000**, *327*, 87-105.

[12] a) B. G. Anderson, W. E. Bauta, J. W. R. Cantrell, T. Engles, D. P. Lovett, *Organic Process Research & Development* **2008**, *12*, 1229-1237; b) Y. Cen, A. A. Sauve, *Nucleosides, Nucleotides and Nucleic Acids* **2010**, *29*, 113-122; c) T. Tennilä, E. Azhayeva, J. Vepsäläinen, R. Laatikainen, A. Azhayev, I. A. Mikhailopulo, *Nucleosides, Nucleotides and Nucleic Acids* **2000**, *19*, 1861-1884.

[13] I. A. Mikhailopulo, *Curr. Org. Chem.* **2007**, *11*, 317-335.

[14] I. A. Mikhailopulo, A. I. Miroshnikov, *Mendeleev Commun.* **2011**, *21*, 57-68.

[15] a) I. A. Mikhailopulo, A. I. Zinchenko, Z. Kazimierczuk, V. N. Barai, S. B. Bokut, E. N. Kalinichenko, *Nucleosides and Nucleotides* **1993**, *12*, 417-422; b) V. N. Barai, A. I. Zinchenko, L. A. Eroshevskaya, E. N. Kalinichenko, T. I. Kulak, I. A. Mikhailopulo, *Helv. Chim. Acta* **2002**, *85*, 1901-1908; c) H. Komatsu, T. Araki, *Nucleosides, Nucleotides and Nucleic Acids* **2005**, *24*, 1127-1130; d) G. Zuffi, S. Monciardini (Explora Laboratories SA, Switz.), *European Patent* EP1932918 A1, **2008**; e) T. A. Krenitsky, G. B. Elion, J. E. Rideout (The wellcome foundation limited), *European Patent* EP0002192 B1, **1981**; f) J. V. Tuttle, M. Tisdale, T. A. Krenitsky, *J Med Chem* **1993**, *36*, 119-125.

[16] S. Taran, K. Verevkina, S. Feofanov, A. Miroshnikov, *Russ. J. Bioorg. Chem.* **2009**, *35*, 739-745.

[17] R. Montilla Arevalo, V. M. Deroncelé Thomas, C. López Gómez, M. Pascual Gilabert, C. Estévez Company, J. Castells Boliart (Institut Univ. de Ciència i Tecnologia), *European Patent* EP 2338985 A1, **2011**

[18] T. A. Krenitsky, G. W. Koszalka, J. V. Tuttle, *Biochemistry-Us* **1981**, *20*, 3615-3621.

[19] D. Ubiali, C. D. Serra, I. Serra, C. F. Morelli, M. Terreni, A. M. Albertini, P. Manitto, G. Speranza, *Adv. Synth. Catal.* **2012**, *354*, 96-104.

[20] a) D. F. Visser, F. Hennessy, J. Rashamuse, B. Pletschke, D. Brady, *J Mol Catal B-Enzym* **2011**, *68*, 279-285; b) D. V. Chebotaev, L. B. Gul'ko, V. P. Veiko, *Russ. J. Bioorg. Chem.* **2001**, *27*, 160-166.

[21] R. Aono, K. Aibe, A. Inoue, K. Horikoshi, *Agr Biol Chem Tokyo* **1991**, *55*, 1935-1938.

[22] a) X. Zhou, K. Szeker, B. Janocha, T. Böhme, D. Albrecht, I. A. Mikhailopulo, P. Neubauer, *FEBS Journal* **2013**, *280*, 1475-1490; b) K. Szeker, X. Zhou, T. Schwab, A. Casanueva, D. Cowan, I. A. Mikhailopulo, P. Neubauer, *J. Mol. Catal. B: Enzym.* **2012**, *84*, 27-34.

[23] a) A. I. Zinchenko, V. N. Barai, S. B. Bokut, E. I. Kvasyuk, I. A. Mikhailopulo, *Appl Microbiol Biotechnol* **1990**, *32*, 658-661; b) V. B. Berzin, E. V. Dorofeeva, V. N. Leonov, A. I. Miroshnikov, *Russ. J. Bioorg. Chem.* **2009**, *35*, 193-196.

Enzymatic Synthesis of 2,6-Dihalogenated Purine Nucleosides

Xinrui Zhou,[a] Kathleen Szeker,[a] Linyu Jiao,[b] Igor A. Mikhailopulo,[c,*] and Peter Neubauer[a,*]

Supporting Information

Contents

1. General 2

2. HPLC Chromatograms 2

3. Purification and Identification of 9 12

4. NMR and HR-MS Spectrum of 9 13

1. General

All chemicals and solvents were of analytical grade or higher and purchased, if not stated otherwise, from Sigma-Aldrich (Steinheim, Germany), Carl Roth (Karlsruhe, Germany), TCI Deutschland (Eschborn, Germany), and VWR (Darmstadt, Germany). Water was purified by a purification system from Merck Millipore (Schwalbach, Germany). MagReSyn^{TM} Epoxide is a kind gift from ReSyn Biosciences (Pretoria, South Africa). The UV spectrum and absorption of purine bases were recorded by Biochrom Ultrospec 3300 pro UV/Visible Spectrophotometer (Cambridge, England). HPLC analysis were carried out with an Agilent 1200 series system equipped with an Agilent DAD detector using a Phenomenex (Torrance, United States) reversed phase C18 column (150 × 4.60 mm). Column chromatography was performed on silica gel 60 (0.043-0.06 mm, 230-400 mesh, ASTM) from Grace GmbH. NMR data were recorded on Bruker AV 500 instruments. High resolution mass spectrometry (HRMS) analyses were performed by the Analytical Facility at the Institut für Chemie of the Technische Universität Berlin.

2. HPLC Chromatograms

HPLC analyses were performed at a flow rate of 1 mL min^{-1}, and the column temperature was 15 °C. The injection volume was 10 mL and followed by a gradient elution: mixture of 20 mM ammonium acetate + 100 % acetonitrile in the ratio of 97:3 to 60:40 in 10 min, and then 97:3 for 8 min as eluent; UV detection at λ=260 nm.

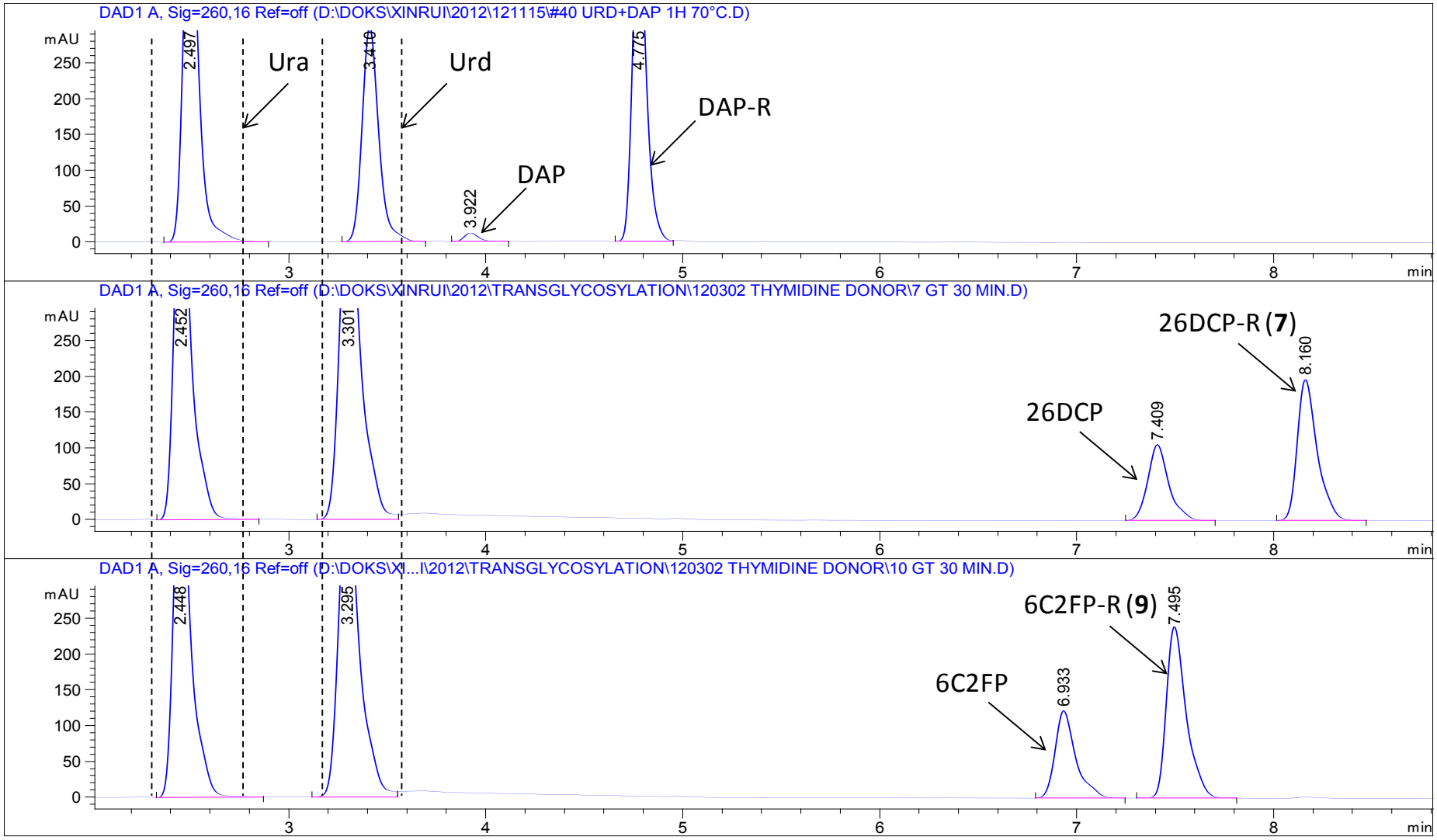


Figure S1. Enzymatic synthesis of DAP-R, 26DCP-R (**7**), and 6C2FP-R (**9**) with Urd as pentofuranosyl donor (entries 1, 4, and 5 refer to Table 3 of the original article).

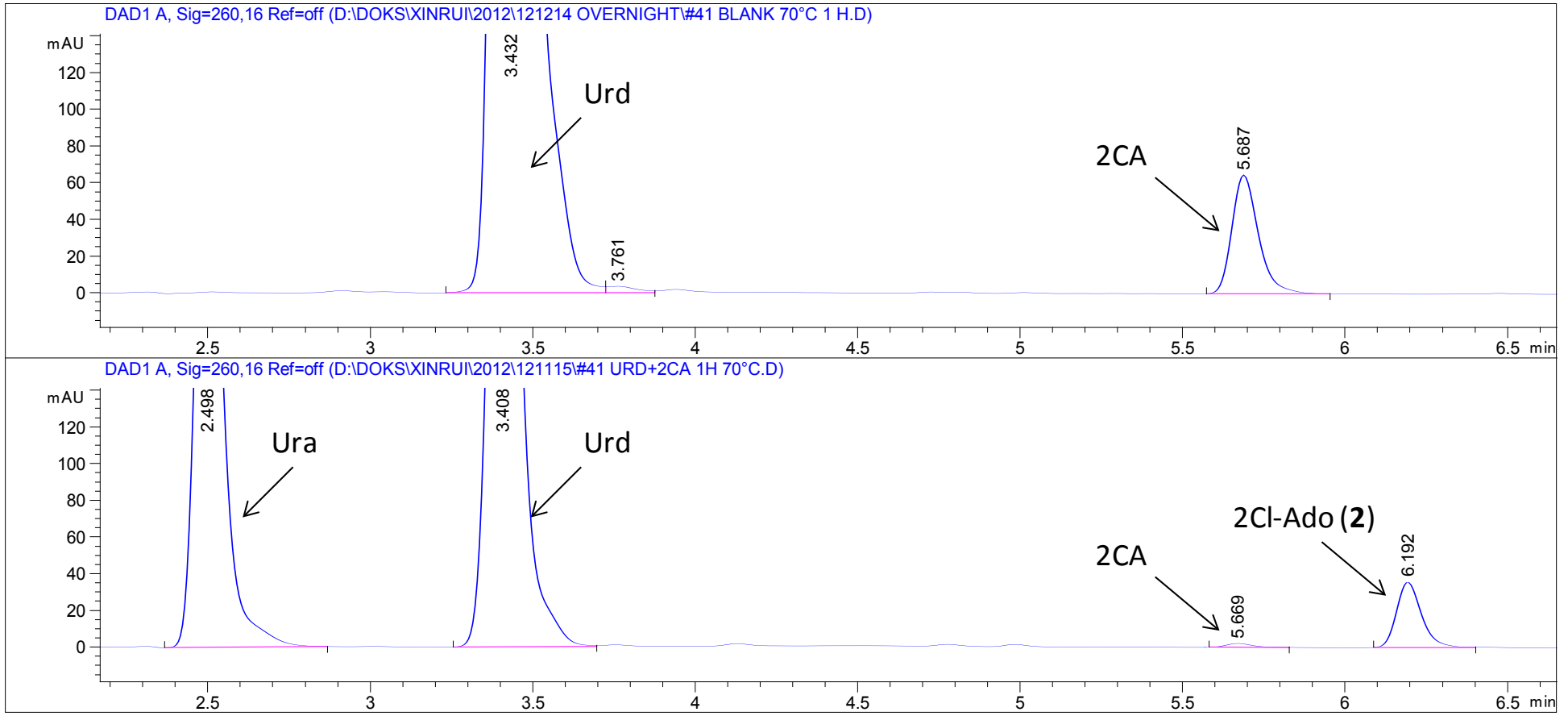


Figure S2. Enzymatic synthesis of 2Cl-Ado (**2**) and the related control (entry 2 refers to Table 3 of the original article).

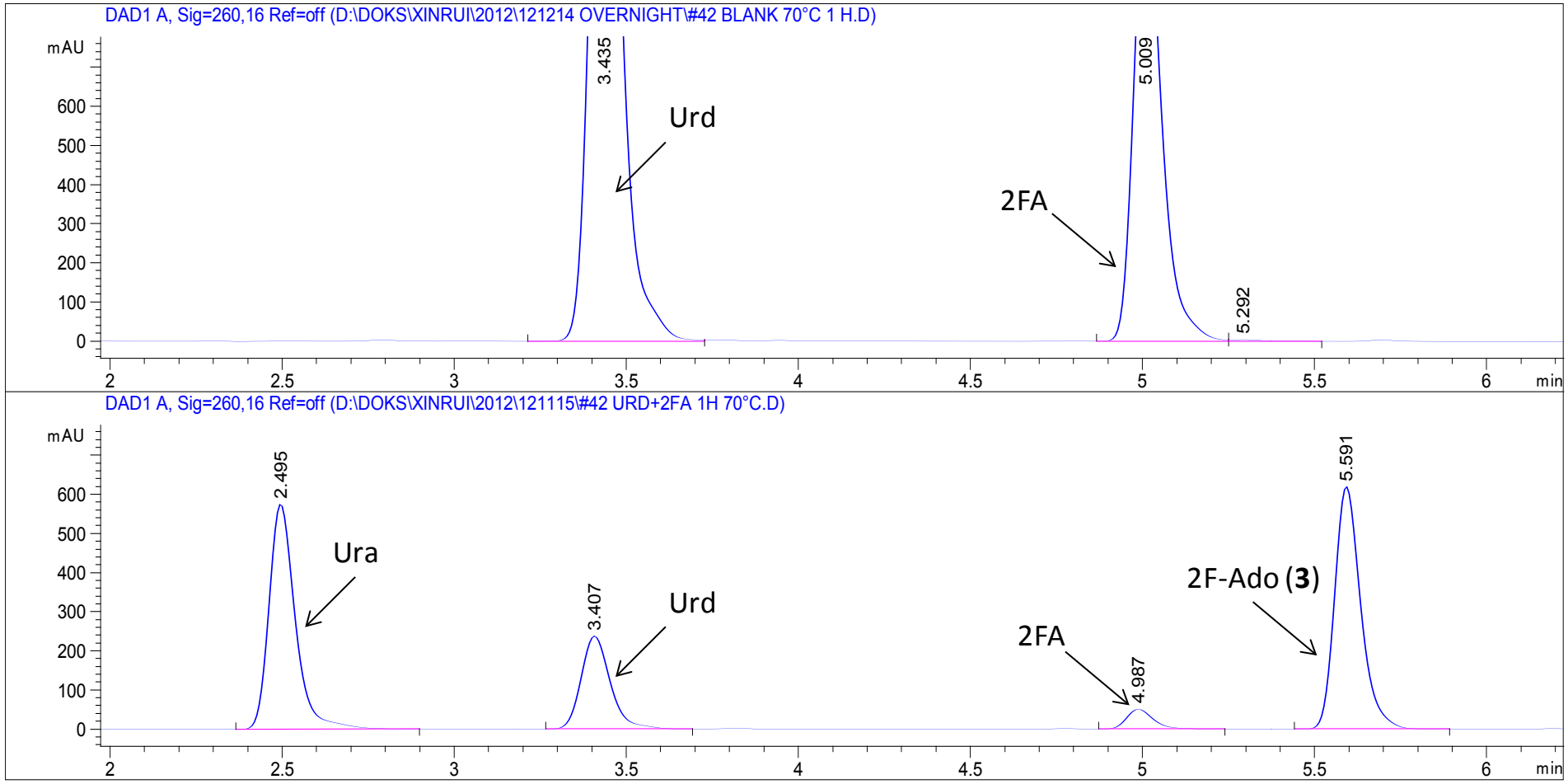


Figure S3. Enzymatic synthesis of 2F-Ado (**3**) and the related control (entry 3 refers to Table 3 of the original article).

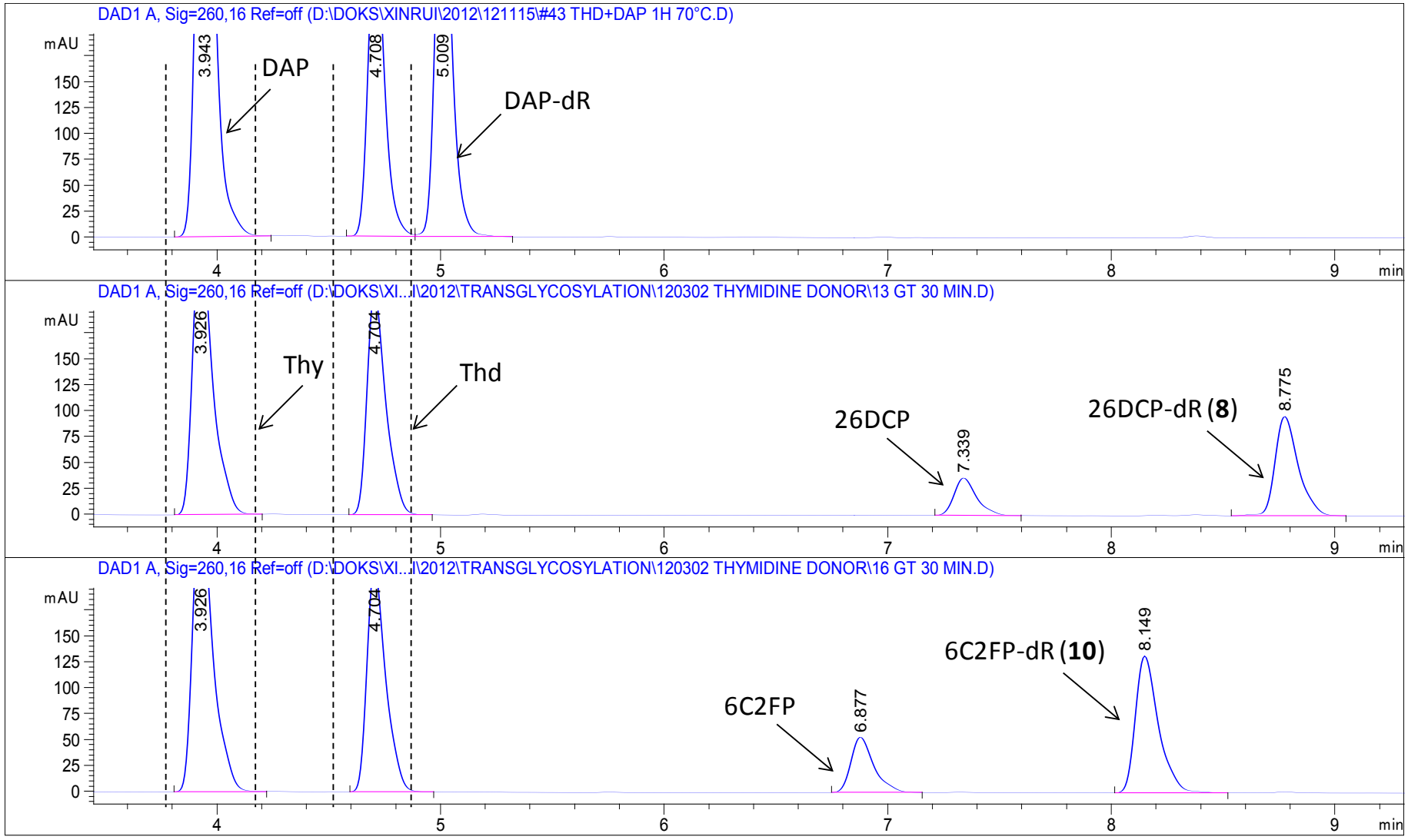


Figure S4. Enzymatic synthesis of DAP-dR, 26DCP-dR (**8**), and 6C2FP-R (**10**) with Thd as pentofuranosyl donor (entries 6, 9, and 10 refer to Table 3 of the original article).

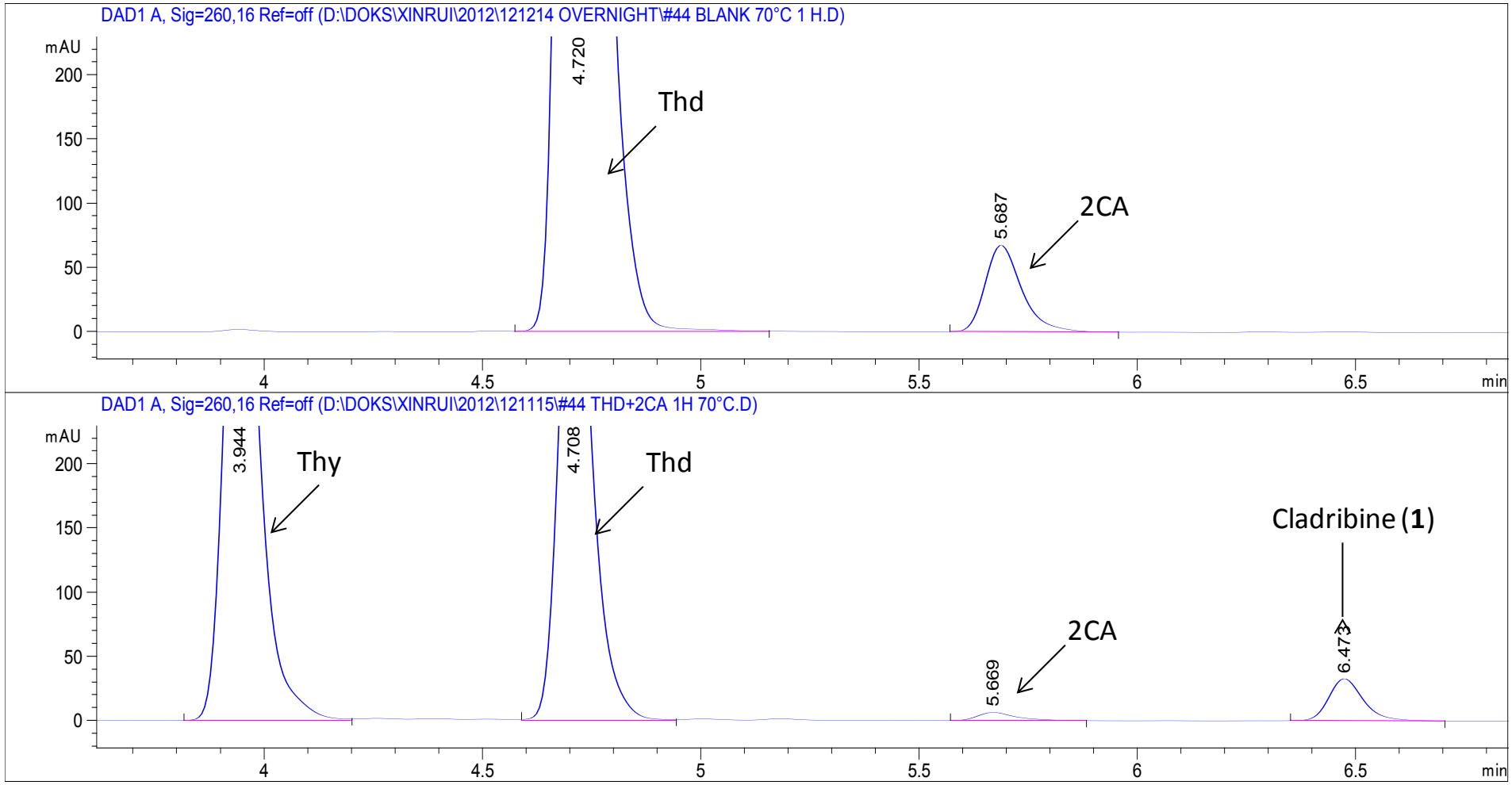


Figure S5. Enzymatic synthesis of Cladribine (**1**) and the related control (entry 7 refers to Table 3 of the original article).

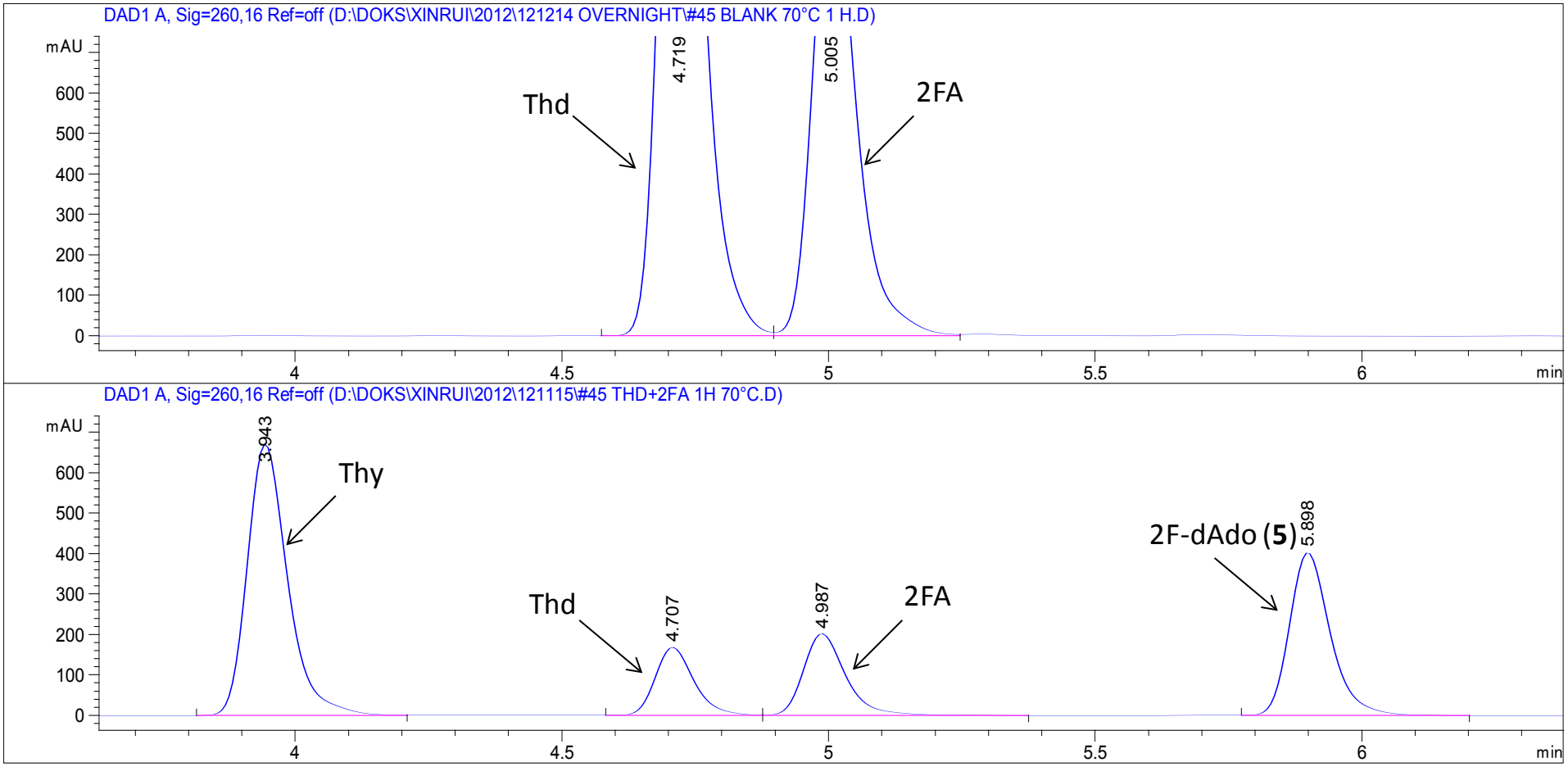


Figure S6. Enzymatic synthesis of 2F-dAdo (**5**) and the related control (entry 8 refers to Table 3 of the original article).

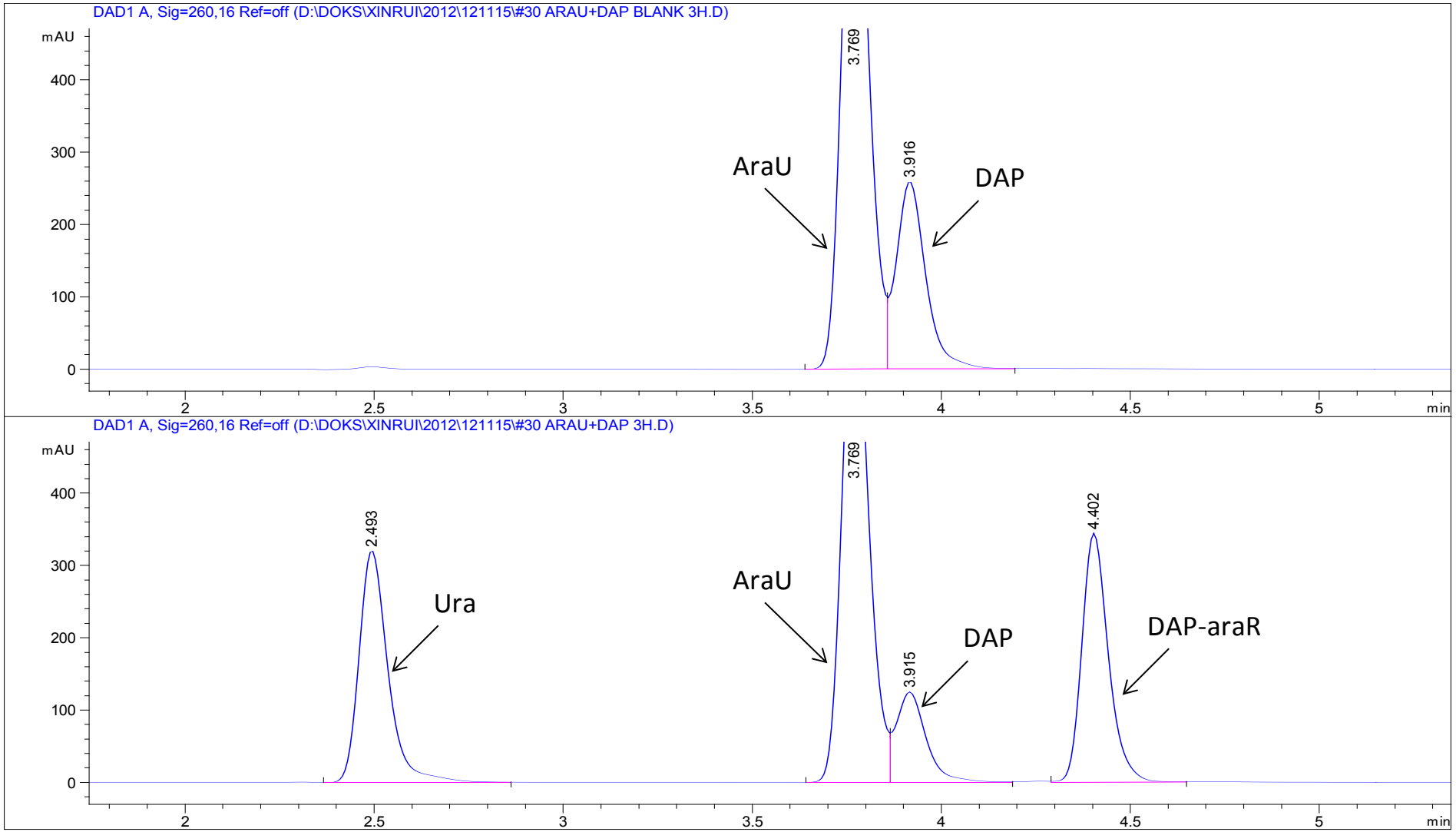


Figure S7. Enzymatic synthesis of DAP and the related control (entry 11 refers to Table 3 of the original article).

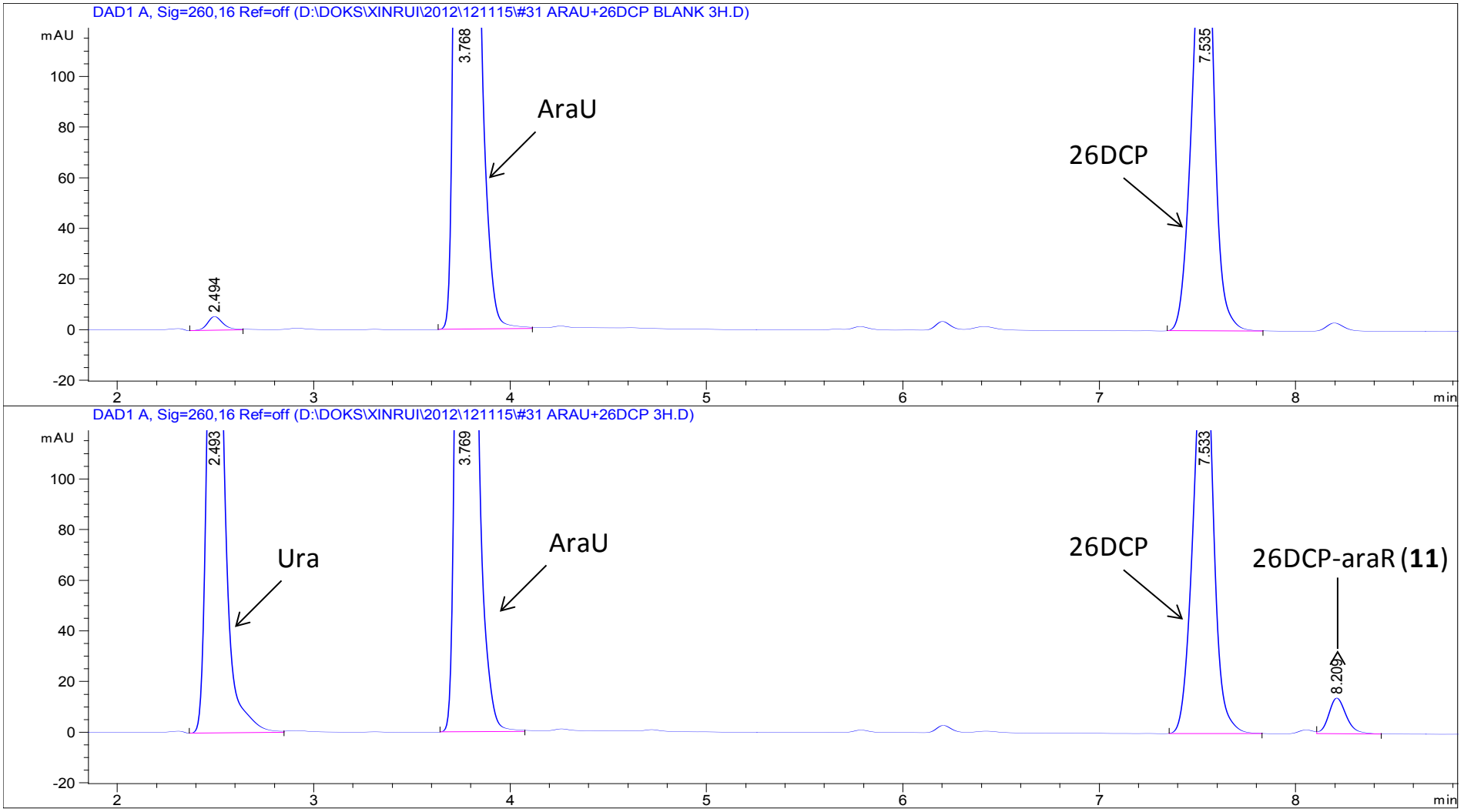


Figure S8. Enzymatic synthesis of 26DCP-araR (**11**) and the related control (entry 12 refers to Table 3 of the original article).

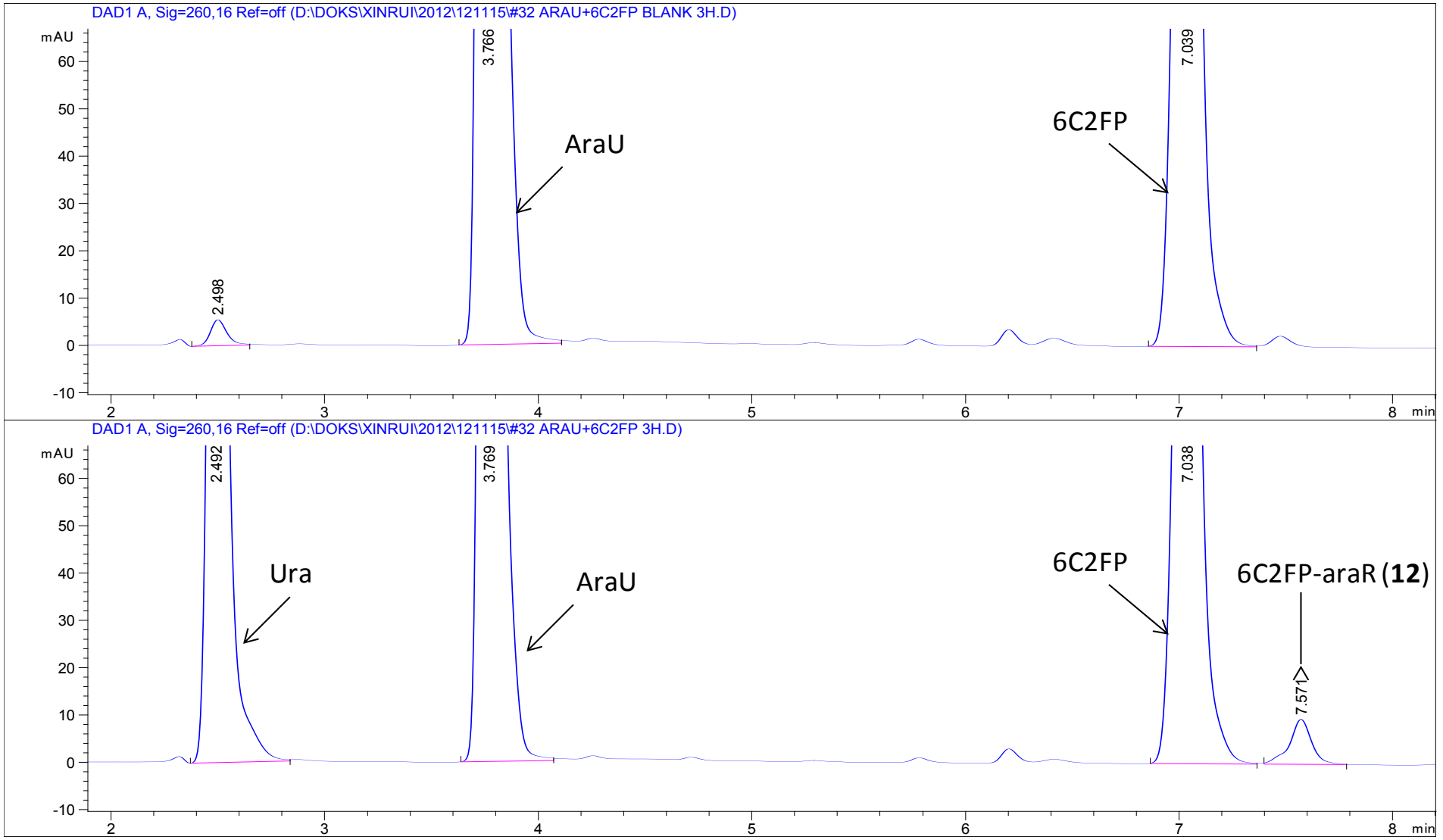


Figure S9. Enzymatic synthesis of 6C2FP-araR (**12**) and the related control (entry 13 refers to Table 3 of the original article).

3. Purification and Identification of 9

After the conversion not increasing with the time any more, the reaction mixture were concentrated under vacuum. The residues were purified on a silica gel column chromatography (dichloromethane: ethanol= 15:1). However, **9** and uracil were not separated successfully and the related fractions were concentrated and subjected to a second silica gel purification (ethyl acetate: acetone= 7:3). Finally the desired product was purified and obtained as a white solid.

Melting point= 182–183 °C (1st run), 182–183.5 °C (2nd run)

Chemical shifts of ^{1}H and ^{13}C NMR are reported in parts per million (ppm) downfield from tetramethylsilane and are referenced to the residual solvent resonance as the internal standard (DMSO: δ = 2.54 ppm for ^{1}H and DMSO-d_6: δ = 40.45 ppm for ^{13}C). Data are reported as follows: chemical shift, multiplicity (br s= broad singlet, s= singlet, d= doublet, t= triplet, dt= doublet of triplets, q= quartet, m= multiplet), coupling constants (Hz) and integration.

^{1}H NMR (500 MHz, DMSO-d_6):
δ= 8.99 (s, 1H; H-8), 5.98 (d, 1H, $J_{1',2'}$ = 4.9 Hz; H-1'), 5.63 (d, J = 5.7 Hz, 1H; 2'-O**H**), 5.29 (d, J = 5.5 Hz, 1H; 3'-O**H**), 5.11 (t, J = 5.5 Hz, 1H; 5'-O**H**), 4.54 (m, 1H, H-2'), 4.22 (m, 1H; H-3'), 4.03 (dt, 1H, $J_{4',3'}$= 4.03 Hz, $J_{4',5'}$= $J_{4',5''}$ = 4.01 Hz; H-4'), 3.74 (ddd, ^{gem}J = 12.0 Hz, J = 5.2 Hz, J = 4.1 Hz, 1H), 3.63 (ddd, ^{gem}J = 11.9 Hz, J = 5.1 Hz, J = 4.0 Hz, 1H) ppm.

^{13}C NMR (126 MHz, DMSO-d_6):
δ= 157.08 (d, J = 213.96 Hz; C-2), 154.48 (d, J = 19.5 Hz; C-6(4)), 151.49 (d, J = 18.2 Hz; C-4(6)), 147.27 (d, J = 2.51 Hz; C-8 through 5 bonds), 131.4 (d, J = 4.8 Hz; C-5), 89.2 (C-1), 86.6 (C-4) 74.9 (C-2), 70.8 (C-3), 61.7 (C-5) ppm.

HR-MS (APCI) exact mass *m/z*= 305.0450, calcd. for $C_{10}H_{10}ClFN_4O_4$ $[M+H]^+$: 305.0447.
Isotope: calcd. for $[M+H]^+$ *m/z*= 307.0418, found: 307.0420.

4. NMR and HR-MS Spectrum of 9

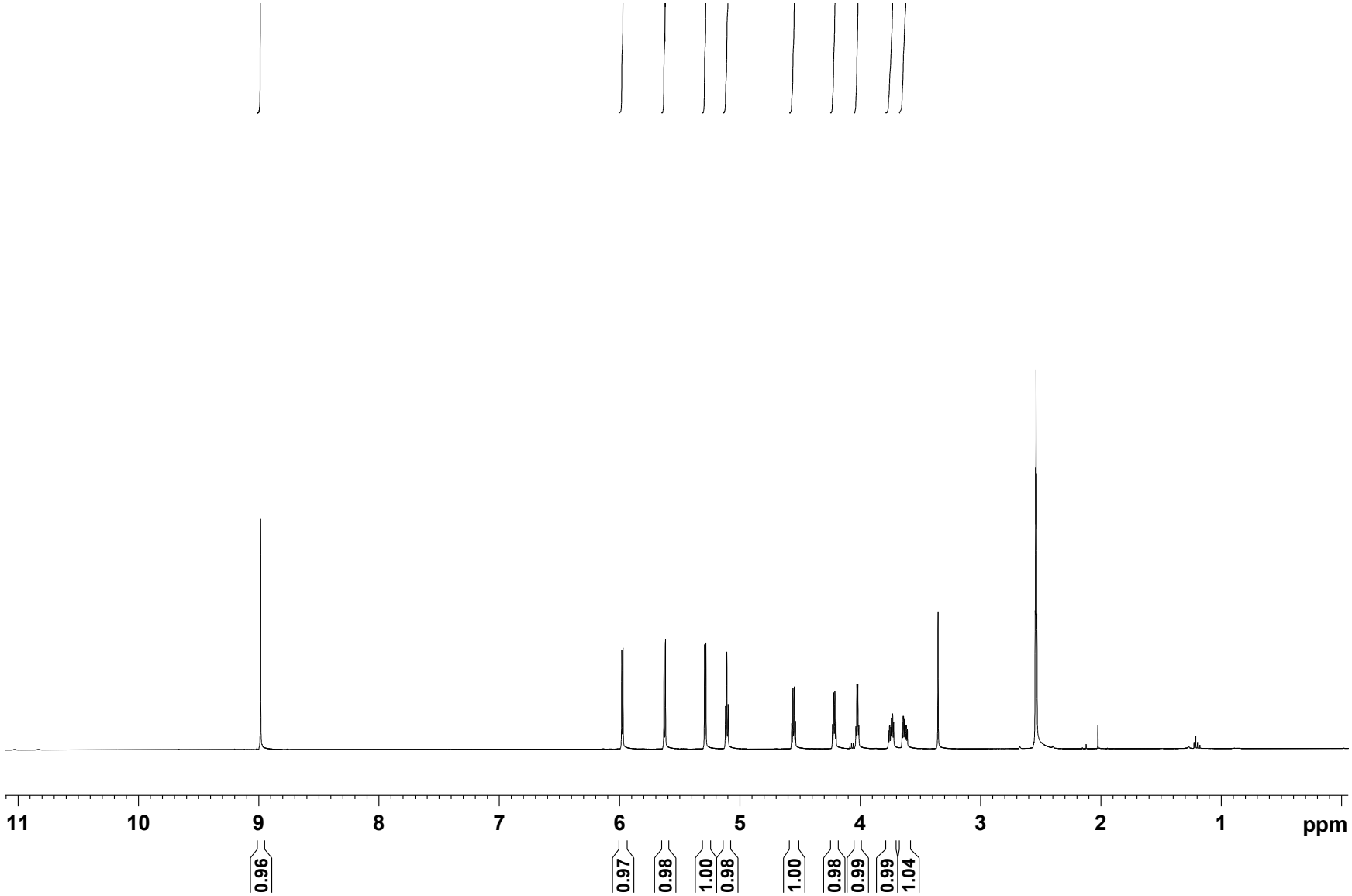


Figure S10. ^{1}H NMR spectrum of **9**

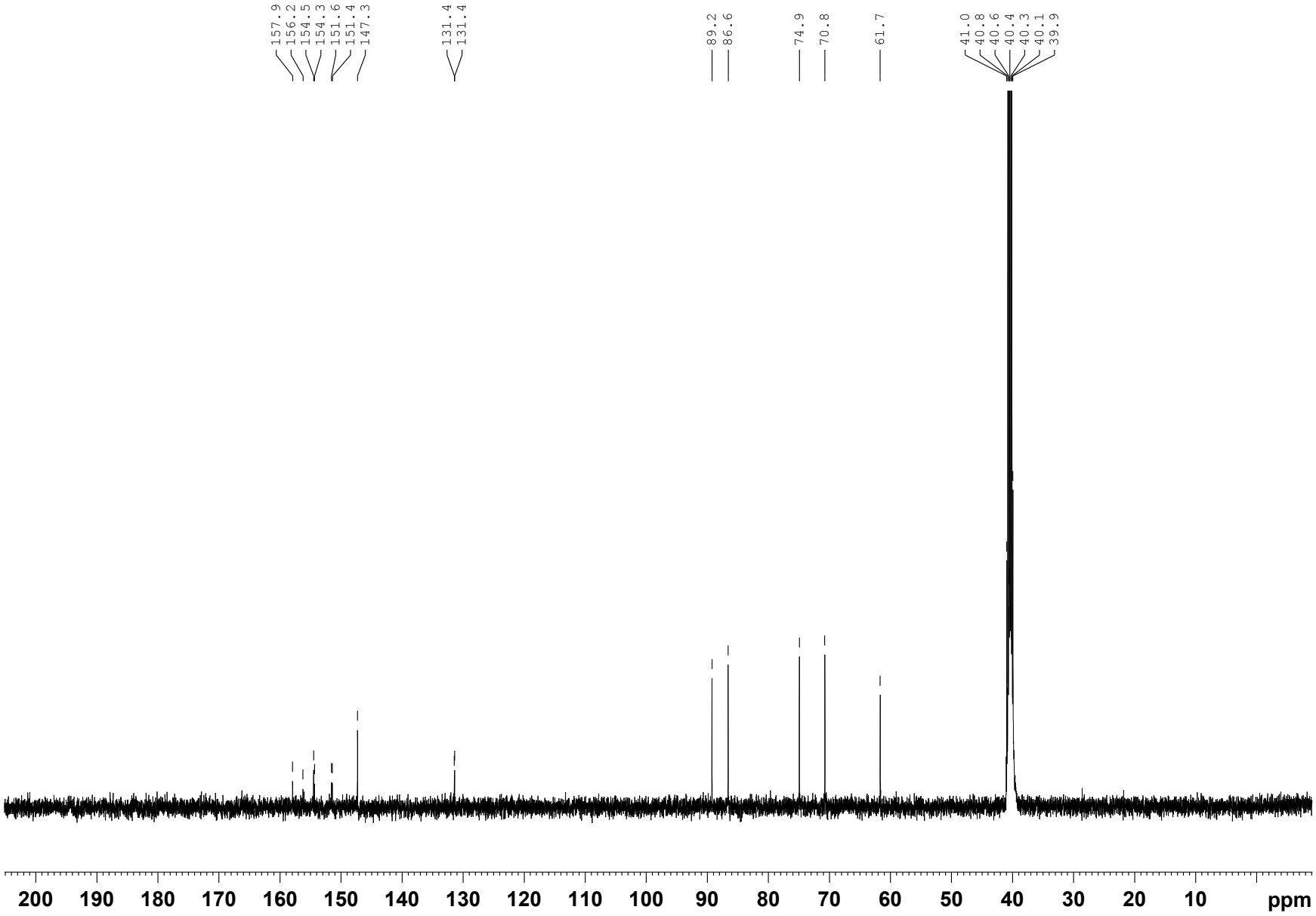


Figure S11. ^{13}C NMR spectrum of **9**

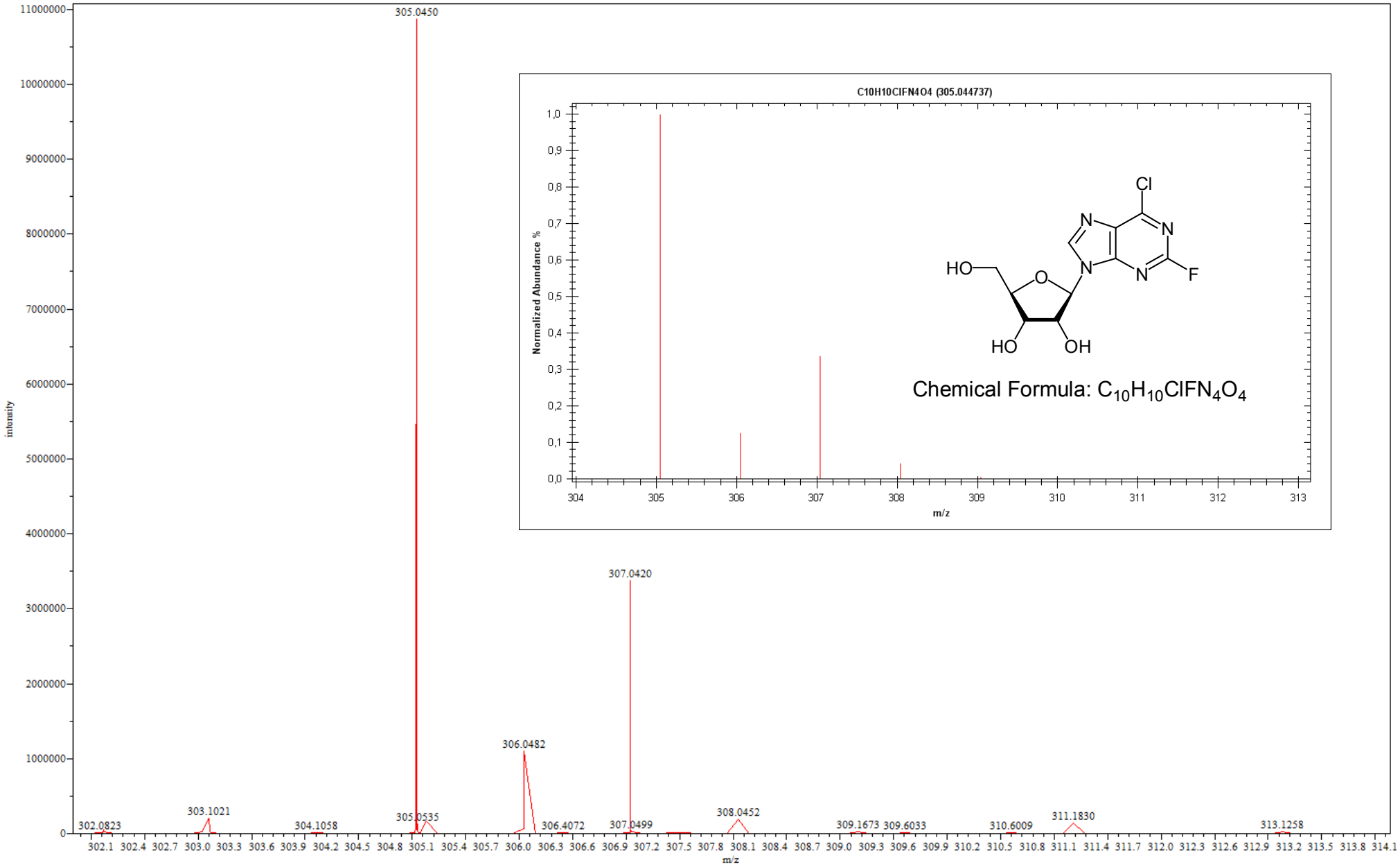


Figure S12. HR-MS Spectrum of **9**. Inserted curve is the calculated isotope distribution of $[M+H]^+$.