

INVESTIGATING STRUCTURAL INSIGHTS AND INHIBITOR STRATEGIES IN MEP PATHWAY ENZYMES FOR DRUG DEVELOPMENT

Dissertation

Zur Erlangung des Grades der Doktorin der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

vorgelegt von

M.Sc. Rawia Abdulazeem Ahmed Hamid

Saarbrücken

2024

Tag des Kolloquiums: 8. Februar 2024

Dekan: Prof. Dr. Ludger Santen

Berichterstatter: Prof. Dr. Anna K. H. Hirsch

Prof. Dr. Jesko-Alexander Köhnke

Akad. Mitglied: Dr. Britta Diesel

Vorsitz: Prof. Dr. Alexandra K. Kiemer

Die vorliegende Arbeit wurde von November 2019 bis Februar 2024 unter Anleitung von Frau Prof. Dr. Anna K. H. Hirsch in der Fachrichtung Medizinische Chemie der Naturwissenschaftlich Technischen Fakultät der Universität des Saarlandes, sowie am Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) in der Abteilung Drug Design and Optimization (DDOP) angefertigt.

Table of contents

Table of contents	I
Acknowledgements.....	III
Thesis Abstract.....	VII
Publications.....	XI
Abbreviations.....	XIII
1 INTRODUCTION	1
1.1 The 2C-Methyl-D-erythritol 4-phosphate (MEP) pathway.....	1
1.2 1-Deoxy-D-xylulose 5-phosphate synthase (DXPS).....	3
1.2.1 Structure of 1-deoxy-D-xylulose 5-phosphate synthase.....	3
1.2.2 Unique catalysis of 1-deoxy-D-xylulose 5-phosphate synthase.....	4
1.2.3 Targeting 1-deoxy-D-xylulose 5-phosphate synthase.....	5
1.2.4 The challenge of DXPS crystallization:	6
1.3 4-Diphosphocytidyl-2-C-methyl-D-erythritol (IspE) kinase.....	7
1.3.1 Structure of the 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase.....	7
1.3.2 Targeting 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase.....	9
1.3.3 Co-crystallizing identified IspE kinase inhibitors is a challenging task.....	10
1.4 Aim of the thesis	11
1.5 References	12
2 RESULTS AND DISCUSSIONS.....	19
2.1 1-Deoxy-D-xylulose-5-phosphate synthase from <i>Pseudomonas aeruginosa</i> and <i>Klebsiella pneumoniae</i> reveals conformational changes upon cofactor binding	20
2.1.1 Supplementary information.....	31
2.2 Design of thiamine analogues for inhibition of thiamine diphosphate (ThDP)-dependent enzymes: Systematic investigation through Scaffold-Hopping and C2-Functionalisation	41
2.2.1 Supplementary information.....	49
2.3 Allosteric Inhibition of DXPS: A New Approach in Drug Discovery.....	148
2.4 IspE Kinase Homologues: Characterization and crystal Structures enable structure-based design of novel inhibitors.....	154
2.4.1 Supplementary information.....	176
3 CONCLUSIONS AND OUTLOOKS	200

Acknowledgements

I extend my deepest appreciation to Prof. *Anna Hirsch*, my primary supervisor, for her steadfast support throughout my academic journey. *Anna* not only guided me through the intricacies of my research but also provided invaluable personal support. Her belief in my potential served as a constant source of motivation, and for that, I am truly grateful.

I am indebted to all my wonderful friends and colleagues who fostered a friendly and collaborative environment during my PhD. The camaraderie we shared made challenging moments more manageable and successes all the more enjoyable. Special thanks to *Erik Kosche*, a great friend who was always there for me in the highs and lows. *Sebastain Adam*, thank you for introducing me to the captivating world of protein crystallography. *Robin Gierse*, your patience in addressing my constant barrage of questions during the early stages of my PhD is deeply appreciated. *Mostafa Hamed* and *Eleonora Diamanti*, my office mates, thank you for creating a great office environment and for tolerating my silly comments and questions.

A heartfelt note of gratitude goes to my co-supervisor, Prof. *Jesko Köhnke*, whose invaluable help and mentorship significantly contributed to the success of my research. His guidance and expertise were instrumental in navigating the complexities of my work, and I am deeply appreciative of his substantial contributions.

Last but certainly not least, I must acknowledge my family—my mother, *Zainab*, and my sister, *Rihab*, back in Sudan. Despite the hardships they endure, their love and encouragement have been my rock. Their unwavering support has fueled my determination, and I am forever grateful for their presence in my life.

Rawia

-Thank you God, for giving me the strength to keep going.

Thesis Abstract

Drug-resistant infections pose a significant threat to human health with studies predicting they will become the leading cause of death in the next 30 years. There is therefore an urgent need to find new antimicrobial agents. Unfortunately, the antimicrobial drug-discovery and development pipeline is sparse, obscuring hopes of finding new drugs in the near future. The methylerythritol (MEP) pathway used mainly by pathogenic microorganisms has emerged as a source of promising new drug targets for the discovery of anti-infective agents, yet it remains underexplored. This thesis presents a comprehensive study of MEP pathway enzymes with a focus on 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) and 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from selected pathogens. We obtained, for the first time, crystal structures of DXPS from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* and explored their structural differences including conformational changes upon binding to the co-factor thiamine diphosphate (ThDP). These findings were instrumental in understanding how the enzyme works and how it could be inhibited. We went on to explore potential inhibition strategies against these enzymes and profiled novel classes of inhibitors, the first was thiamine monophosphate (ThMP) analogues as ThDP competitive inhibitors of DXPS, supported by binding mode analyses by obtaining a co-crystal structure with *K. pneumoniae* DXPS. Second, we present a promising breakthrough in DXPS inhibition by targeting an allosteric binding site and provide experimental evidence in the form of a co-crystal structure of an allosteric inhibitor with *P. aeruginosa* DXPS. Furthermore, we analyzed the kinetics and structural variances among IspE kinase homologs. We explored the potential impact of these structural differences on the binding mode of inhibitors, substantiating our findings through co-crystal structures with various inhibitors. This approach offers valuable insights into potential strategies for designing inhibitors. Overall, this work presents novel findings with promising potential to guide antimicrobial drug discovery targeting drug-resistant pathogens and employing underexplored enzyme inhibition approaches.

Zusammenfassung

Arzneimittelresistente Infektionen stellen eine erhebliche Bedrohung für die menschliche Gesundheit dar. Studien prognostizieren, dass sie in den nächsten 30 Jahren die führende Todesursache werden könnten. Daher besteht ein dringender Bedarf, neue antimikrobielle Wirkstoffe zu entdecken. Leider ist die Pipeline für die Entdeckung und Entwicklung antimikrobieller Medikamente dünn gesät, was die Hoffnung auf die baldige Entdeckung neuer Medikamente trübt. Der Methylerythritol (MEP)-Weg, der hauptsächlich von pathogenen Mikroorganismen genutzt wird, hat sich als vielversprechende Quelle neuer Wirkstoffziele für die Entdeckung von antiinfektiven Substanzen erwiesen, bleibt jedoch unterexploriert. Diese Dissertation präsentiert eine umfassende Studie zu Enzymen des MEP-Wegs mit Schwerpunkt auf 1-Deoxy-D-xylulose-5-phosphat-Synthase (DXPS) und 4-Diphosphocytidyl-2-C-methyl-D-erythritol-Kinase (IspE) aus ausgewählten Pathogenen. Zum ersten Mal erhielten wir Kristallstrukturen von DXPS aus *Pseudomonas aeruginosa* und *Klebsiella pneumoniae* und erforschten ihre strukturellen Unterschiede, einschließlich konformationeller Veränderungen bei der Bindung an den Cofaktor Thiamindiphosphat (ThDP). Diese Erkenntnisse waren entscheidend, um zu verstehen, wie das Enzym funktioniert und wie es gehemmt werden könnte. Wir untersuchten dann potenzielle Hemmstrategien gegen diese Enzyme und profilten neuartige Klassen von Inhibitoren. Die erste umfasste Thiaminmonophosphat (ThMP)-Analoge als ThDP-konkurrenzfähige Inhibitoren von DXPS, unterstützt durch Bindungsmodusanalysen durch Erlangen einer Co-Kristallstruktur mit *K. pneumoniae* DXPS. Zweitens präsentieren wir einen vielversprechenden Durchbruch in der Hemmung von DXPS durch die Ausrichtung auf eine allosterische Bindungsstelle und liefern experimentelle Beweise in Form einer Co-Kristallstruktur eines allosterischen Inhibitors mit *P. aeruginosa* DXPS. Darüber hinaus analysierten wir die Kinetik und strukturellen Unterschiede zwischen IspE-Kinase-Homologen. Wir erforschten die potenziellen Auswirkungen dieser strukturellen Unterschiede auf den Bindungsmodus von Inhibitoren und untermauerten unsere Erkenntnisse durch Co-Kristallstrukturen mit verschiedenen Inhibitoren. Dieser Ansatz bietet wertvolle Einblicke in potenzielle Strategien zur Gestaltung von Inhibitoren. Insgesamt präsentiert diese Arbeit neuartige Erkenntnisse mit vielversprechendem Potenzial, die antimikrobielle Wirkstoffentwicklung zur Bekämpfung von resistenten Pathogenen zu lenken und unterexplorierte Ansätze zur Enzymhemmung zu nutzen.

Publications

Publications included in this thesis

Rawia Hamid, Sebastian Adam, Antoine Lacour, Leticia Monjas Gomez, Anna K. H. Hirsch. 1-deoxy-D-xylulose 5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding. *J Biol Chem.* (2023). doi: 10.1016/j.jbc.2023.105152

Chan, Alex HY, Terence CS Ho, Rimsha Irfan, Rawia Hamid, Emma S. Rudge, Amjid Iqbal, Alex Turner, Anna KH Hirsch, and Finian J. Leeper. "Design of thiamine analogues for inhibition of thiamine diphosphate (ThDP)-dependent enzymes: Systematic investigation through Scaffold-Hopping and C2-Functionalisation." *Bioorganic Chemistry* 138 (2023): 106602. doi: 10.1016/j.bioorg.2023.106602.

Rawia Hamid, Danica J. Walsh, Eleonora Diamanti, Diana Aguilar, Antoine Lacour, Mostafa Hamed, Anna K. H. Hirsch. IspE Kinase Homologues: Characterization and crystal Structures enable structure-based design of novel anti-infectives. *Submitted manuscript.*

Publications not included in this thesis:

Gierse, Robin M., Eswar R. Reddem, Alaa Alhayek, Dominik Baitinger, Rawia Hamid, Harald Jakobi, Bernd Laber, Gudrun Lange, Anna KH Hirsch, and Matthew R. Groves. "Identification of a 1-deoxy-D-xylulose-5-phosphate synthase (DXS) mutant with improved crystallographic properties." *Biochemical and Biophysical Research Communications* 539 (2021): 42-47.

Contribution: Protein purification and enzyme assays optimization.

Jumde, Ravindra P., Melissa Guardigni, Robin M. Gierse, Alaa Alhayek, Di Zhu, Rawia Hamid, Sandra Johannsen et al. "Hit-optimization using target-directed dynamic combinatorial chemistry: development of inhibitors of the anti-infective target 1-deoxy-D-xylulose-5-phosphate synthase." *Chemical Science* 12, no. 22 (2021): 7775-7785.

Contribution: Protein purification, enzyme assays and crystal screening.

Johannsen, Sandra, Robin M. Gierse, Aleksandra Olshanova, Ellie Smerznak, Christian Laggner, Lea Eschweiler, Zahra Adeli, Rawia Hamid, Alaa Alhayek, et al. "Not Every Hit-Identification Technique Works on 1-Deoxy-D-Xylulose 5-Phosphate Synthase (DXPS): Making the Most of a Virtual Screening Campaign." *ChemMedChem* (2023): e202200590.

Contribution: Protein purification and enzyme assays.

Ropponen, Henni-Karoliina, Eleonora Diamanti, Sandra Johannsen, Boris Illarionov, Rawia Hamid, Miriam Jaki, Peter Sass, Markus Fischer, Jörg Haupenthal, and Anna Katharina Herta Hirsch. "Exploring the translational gap of a novel class of *Escherichia coli* IspE inhibitors." *ChemMedChem* (2023): e202300346..

Contribution: Protein purification, enzyme assays and crystal screening.

Abbreviations

Listed in their order of appearance in the text:

DXPS = 1-deoxy-D-xylulose-5-phosphate synthase

GAP = glyceraldehyde 3-phosphate

ThDP = thiamine diphosphate,

IspC = 1-deoxy-D-xylulose-5-phosphate reductoisomerase

NADPH = nicotinamide adenine dinucleotide phosphate

IspD = 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase

CTP = cytidine triphosphate

IspE = 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase

CDP-ME = 4-diphosphocytidyl-2-C-methylerythritol

ATP = adenosine triphosphate

ADP = adenosine diphosphate

IspF = 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

CMP = cytidine monophosphate

CDP-MEP = 4-diphosphocytidyl-2-C-methylerythritol

IspG = 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

IspH = 4-hydroxy-3-methylbut-2-enyl diphosphate reductase

MEcDP = 2-C-methyl-D-erythritol 2,4-cyclodiphosphate

HMBDP = 4-hydroxy-3-methyl-but-2-enyl diphosphate

IDP = isopentenyl diphosphate

DMADP = dimethylallyl diphosphate

Chapter 1

Introduction

1 INTRODUCTION

The rise of resistant strains among different types of pathogens has resulted in globally untreatable diseases classified by the WHO (1, 2). There is a growing consensus among scientists and policymakers to adopt fresh scientific and logistic strategies for anti-infective treatment. The 2C-Methyl-D-erythritol 4-phosphate (MEP) pathway was recently discovered and recognized as prospective anti-infective target, due to its role in the production of isoprenoids, especially in pathogenic microorganisms (3–6).

1.1 The 2C-Methyl-D-erythritol 4-phosphate (MEP) pathway

Isoprenoids, a broad category of natural products, include a wide variety of substances, are present in all living organisms and they play an essential role in cellular and biochemical processes, e.g. cell-wall bio-synthesis, signaling, post translational modification of signaling proteins and synthesis of vitamins. Isoprenoids are also integral to plant defense mechanisms. With over 55,000 known members, they constitute a large group of natural substances found in all areas of life (7, 8).

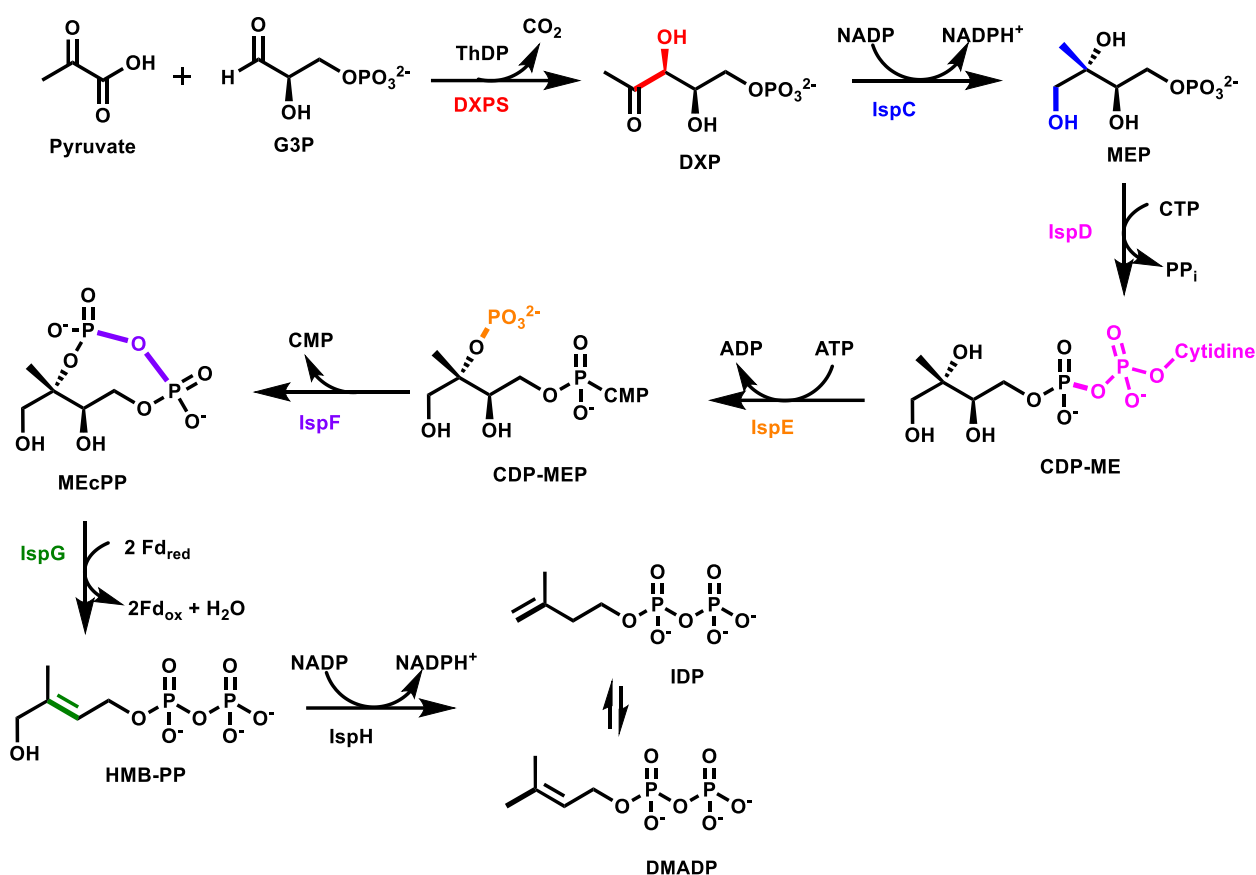
These compounds are produced from either isopentenyl diphosphate (IDP) or dimethylallyl diphosphate (DMADP). Isoprenyl diphosphate synthases are enzymes responsible for combining IDP with allylic primer substrates, biosynthesis involves regulated chain elongation reactions that occur sequentially and stop at specific chain lengths, resulting in the formation of the fundamental linear structures found in all isoprenoid compounds. These compounds can be categorized based on their carbon chain lengths, which include monoterpenes, sesquiterpenes, diterpenes, and triterpenes (9).

The Mevalonate pathway was long considered to be the only source of isoprene units, in spite numerous conflicting observations, particularly in bacteria and plants e.g., the synthesis of mono and diterpenes, and carotenoids (10, 11). However, in the early 1990s, the MEP pathway was discovered as distinct route for the biosynthesis of isoprenoids in plants and bacteria (3).

The MEP pathway is predominantly used by the chloroplasts in phototrophic organisms and is also the primary biosynthetic route for IDP and DMADP in many pathogenic bacteria and malaria parasites (3, 12, 13). Although the majority of microorganisms depend on the MEP pathway for synthesizing these precursors, this pathway is notably lacking in humans. Consequently, the enzymes from this pathway hold great promise as targets for the development of selective inhibitors. The recent understanding of the MEP pathway has unlocked opportunities for devising novel strategies against microbial pathogens and creating specialized inhibitors with broad-spectrum antimicrobial properties (12, 14, 15).

The MEP pathway is a collective route made up of seven enzymes (illustrated in scheme 1 Fig. 1) (16). Of all the enzymes in the MEP pathway, DXPS holds particular significance due to its dual role in catalysis and regulation, it has the highest flux-control coefficient among the

enzymes of the MEP pathway (17). The product DXP not only participates in the MEP pathway but also contributes to the production of thiamine (vitamin B1) and pyridoxine (vitamin B6). Therefore, targeting DXPS for inhibition offers an effective double sided approach compared to other interventions (18, 19). Targeting MEP pathway enzymes by small-molecule inhibitors has been demonstrated through the antibiotic fosmidomycin which primarily inhibits the second step in this pathway and has been applied to treat malaria in mice infected with *Plasmodium vinckei*. Furthermore, fosmidomycin is in clinical trials for evaluation as a treatment for acute *Falciparum* malaria in humans as part of combination therapy (20–23).



Scheme 1 : MEP pathway. The MEP pathway begins with DXPS, which serves as the rate-controlling step. DXP is formed through the condensation of pyruvate and D-GAP. MEP is subsequently produced from DXP in a reduction reaction catalyzed by IspC. Following this, MEP undergoes a sequence of phosphorylation and cyclization reactions, facilitated successively by IspD, IspE, and IspF, resulting in the production of MRcPP. MRcPP is then converted by IspG through a double one-electron transfer reaction into HMB-PP. HMB-PP is further reduced to IDP/DMADP by IspH. Enzymes and substrates full names are shown in the abbreviation section.

1.2 1-Deoxy-D-xylulose 5-phosphate synthase (DXPS)

The gene responsible for encoding DXPS was initially identified in *E. coli* and labeled as *dxs*. The product of the *dxs* gene was found to share sequence similarities with transketolase protein family and later identified as a member of this family. DXP synthase was subsequently recognized as being responsible for catalyzing the production of 1-deoxy-D-xylulose (DXP), which is the first product in the MEP pathway. Furthermore, DXP was acknowledged as a precursor for other essential molecules such as vitamin B1 and B6 (24).

1.2.1 Structure of 1-deoxy-D-xylulose 5-phosphate synthase

DXPS belongs to the thiamine diphosphate (ThDP)-dependent enzyme subfamily within the transketolase family. Like other ThDP-dependent enzymes, DXPS operates as a dimer. In its structural configuration, each monomer within this dimer consists of three linearly arranged domains, although there is a slight bend. Each of these domains is characterized by an α/β fold, with α helices enveloping central, parallel, five to six-stranded β -sheets.

In contrast to other transketolase family members, which position their dimers in a twisted configuration, the DXPS protein forms nearly parallel and interconnected dimers. This means that the domains of one monomer are directly linked to those in the adjacent monomer. This distinction in how the catalytic center is organized, whether within a single molecule or between molecules, serves as a distinguishing feature of ThDP-dependent enzymes. It's essential to note, however, that the ThDP binding site and pocket geometry remain remarkably consistent across the ThDP-dependent enzyme family (25–27).

The ThDP-binding site is positioned at the interface of domains I and domain II, domain II is oriented towards the pyrimidine ring, while domain I faces the diphosphate moiety. ThDP binds in a bent configuration, the pyrimidine ring of ThDP forms a π - π stacking interaction with phenylalanine, N3 makes a direct bonding interaction with the serine/alanine backbone of DXPS. Although ThDP is mostly buried inside the active site, the middle thiazolium ring, particularly position C2, which represent the reactive part of ThDP, is exposed to the solvent, allowing for pyruvate and glyceraldehyde 3-phosphate (D-GAP) required interaction. On the other hand the diphosphate group interacts with several polar residues in domain I and contributes to the sequestering of a magnesium ion (Fig. 1) (28, 29).

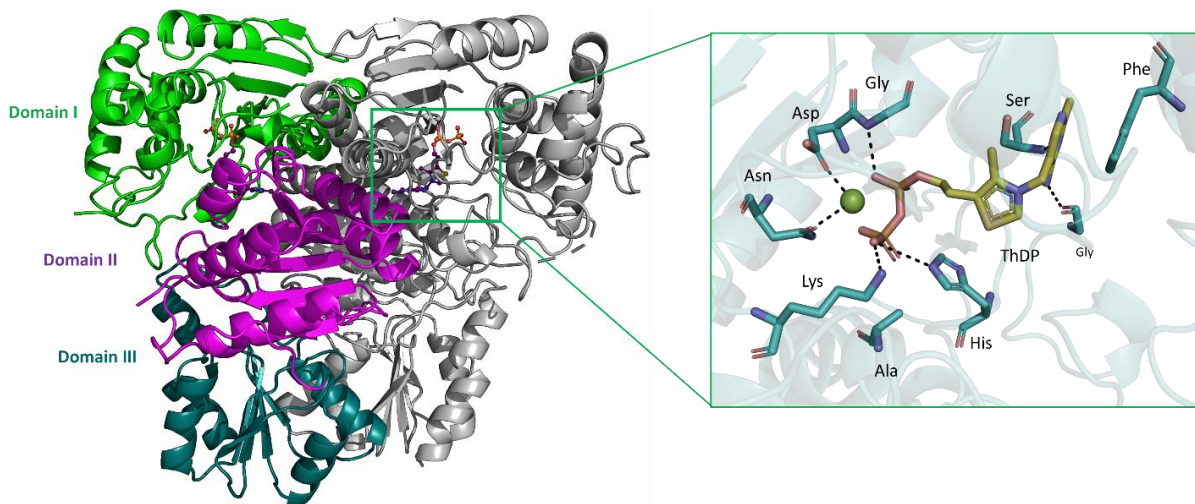
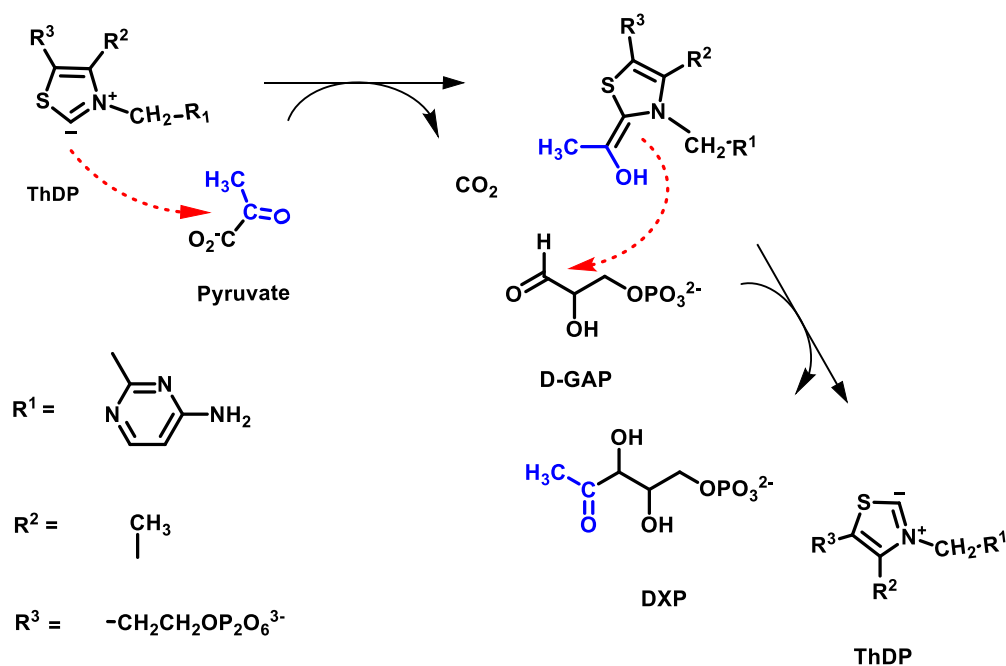


Figure 1: Overview of DXPS from *Escherichia coli*, dimer and domain organization and a zoomed in view of the catalytic center and ThDP binding, positioned between domain I and II, ThDP binds to DXPS mainly through stacking of the thiamine moiety in a hydrophobic pocket and the coordination of the diphosphate moiety and the magnesium (shown as a green sphere) within a network of polar residues.

1.2.2 Unique catalysis of 1-deoxy-D-xylulose 5-phosphate synthase

In 2011, Brammer et al. presented a detailed description of DXPS's catalytic mechanism, introducing a novel random, sequential kinetic model, which is unique within this enzyme category. According to their findings, D-GAP and pyruvate independently and reversibly bind to DXPS synthase (30). This discovery marked a significant departure from the previously proposed ordered or ping-pong mechanisms that had been associated with DXPS catalysis. (4, 31). A subsequent research endeavor uncovered that the enzyme is able to stabilize the ThDP-bound pre-decarboxylation intermediate, specifically the C2 α -lactylThDP (LThDP), even when GAP is absent. The decarboxylation step is dependent on the formation of a tertiary complex where both LThDP and GAP are bound, then the LThDP-DXPS complex reacts with GAP to generate DXP. Subsequent investigations have also revealed that this phenomenon can be induced by the presence of oxygen (O₂) (32, 33). Overview of the reaction catalyzed by DXPS is depicted in scheme 2.

DXPS exhibits an impressive ability to stabilize the catalytic intermediate LThDP within its active site. Mutagenesis investigations conducted on *Deinococcus radiodurans* DXPS have shed light on the specific functions of certain amino acids in binding substrates and intermediates during catalysis. Notably, in domain II, Arg-423, Asp-430, and His-434 have been identified as playing unique roles in stabilizing intermediates, distinct from their counterparts in transketolase. Additionally, these mutagenesis studies have unveiled a potential backup mechanism involving His-82 and His-304 during the catalytic process (34, 35). Subsequent x-ray crystallography studies of DrDXPS-bound intermediates have illustrated how these histidine residues collaborate to support the stability of the catalytic intermediates (36).



Scheme 2. DXPS unique mechanism. DXPS mediates the transfer of an acetaldehyde group from pyruvate to D-GAP, which functions as the acceptor molecule. Production of DXP is accompanied by the release carboxylic group of pyruvate in the form of carbon dioxide (CO_2).

1.2.3 Targeting 1-deoxy-D-xylulose 5-phosphate synthase

DXPS is notably significant due to its dual role in both catalyzing and regulating the MEP pathway. The product, DXP, not only plays a crucial role in the MEP pathway but also contributes to the synthesis of thiamine (vitamin B1) and pyridoxine (vitamin B6). Consequently, targeting DXPS for inhibition presents a highly effective approach with dual benefits when compared to other interventions (18, 19).

Ketoclofazone, a hydrophilic small molecule, derived from the herbicide clomazone, was among the first reported inhibitors of DXPS in both *E. coli* and *Haemophilus influenzae*. Studies were not able to identify its mode of binding, but could show an uncompetitive inhibition against pyruvate, specifically binding to the holoenzyme complex rather than the apo enzyme. In addition, a validation test using DXP, confirmed that ketoclofazone targets the DXPS intracellularly (31, 37).

Earlier DXPS Inhibitors were developed to target the cofactor binding site using ThDP analogues and derivatives (38–40). The use of ThDP analogues as competitive inhibitors lacks high selectivity among different enzyme types. However, DXPS, stands out due to its distinct catalytic behavior and substrate binding site. Targeting this substrate-binding pocket holds promise for achieving selective inhibition. Recently, acylphosphonates have been employed as DXPS inhibitors with high selectivity. These inhibitors effectively engage with DXPS and could potentially extend their reach to the substrates, pyruvate and D-GAP, binding site, resulting in bi-substrate inhibitors achieving selectivity over other ThDP dependent enzymes.

Subsequent optimizations of these inhibitors class have led to the development of enamine prodrugs of acetylphosphonates. The aim was to enhance active uptake of the compound inside bacteria via the oligopeptide permease, resulting in a significant improvement in their *in-vivo* activity against Gram-negative pathogens (41–43). These recent developments suggest a shift from considering alkylphosphonates as mere tools for studying the enzymatic mechanism of DXPS to viewing them as candidates with more drug-like properties (Fig. 2).

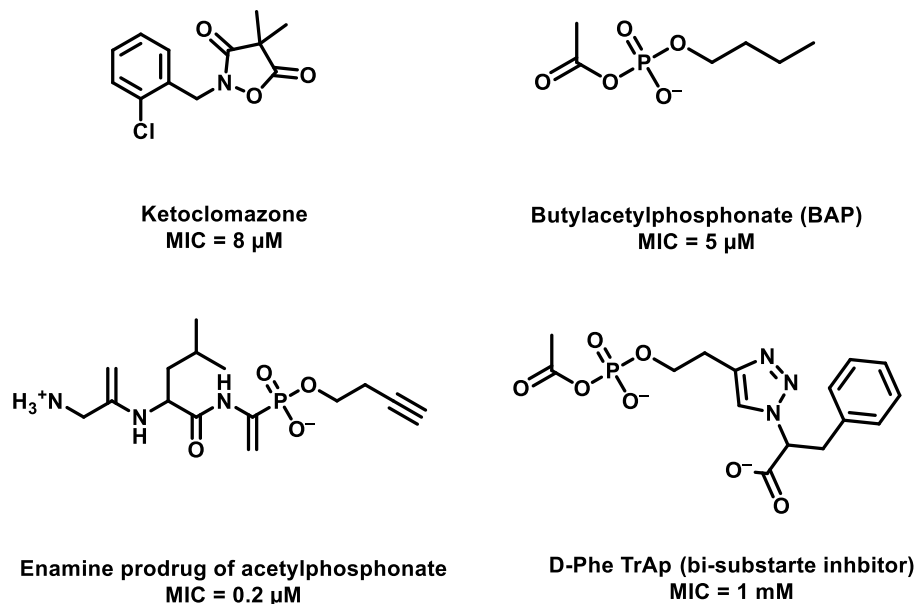


Figure 2: activity values of some of available DXPS inhibitors tested against *Escherichia coli*.

1.2.4 The challenge of DXPS crystallization:

In 2007, the initial structural insights into DXS from *E. coli* and *D. radiodurans*, and its interaction with the cofactor ThDP were obtained. The DXPS structure was resolved at a 2.4 Å resolution, employing the selenomethionyl anomalous diffraction method. Notably, the successful crystallization of DXPS depended on *in situ* proteolysis performed by a fungal protease. It was later observed that two segments of the enzyme, which are the connecting loops between its three domains, did not exhibit any electron density. Analyzing the remaining portions of the enzyme, including their molecular weights and species observed in SDS gels of the protein, confirmed that these two absent segments had been removed by the fungal protease. These segments are indicative of flexible loops in the structure (25).

Further research has indicated that when parts of these flexible segments are eliminated from DXPS in *D. radiodurans*, it maintains its enzymatic activity while reducing protein degradation, and improving crystallization characteristics (Fig. 3) (44). Consequently, this protein engineering method has enabled the successful crystallization of DXPS truncated constructs, facilitating obtaining structural information from pathogenic species that were previously resistant to crystallization (26, 45, 46).



Figure 3: DXPS monomer structure showing the truncated loop. Sequence alignment of the published structures from *Deinococcus radiodurans* and *Mycobacterclosis tuberculosis* and *E. coli* shows low sequence conservation in the loop region. The truncated part is highlighted in red in both structure and sequence.

1.3 4-Diphosphocytidyl-2-C-methyl-D-erythritol (IspE) kinase

IspE kinase is another essential enzyme within the MEP pathway. In 2000, The *YchB* gene in *E. coli* was identified as the responsible gene for the production of IspE. The enzyme was then identified to catalyze the ATP-dependent phosphorylation of its substrate CDP-ME. In addition IspE plays a central role in a three-step reaction that converts MEP to the cyclic diphosphate, MEcDP, alongside the enzymes IspD and IspF (47, 48).

1.3.1 Structure of the 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase

The IspE structure was fully characterized in 2003 through the analysis of crystal structures obtained from two sources: *Thermus thermophilus* and *E. coli*. The enzyme primarily consists of two domains: the N-terminal domain, which contains the ATP cofactor binding site, and the C-terminal domain, responsible for accommodating the substrate, CDP-ME. Both domains exhibit a characteristic α/β fold. The catalytic pocket, notably wide and exposed to the solvent, is positioned between these two domains. This arrangement allows for the close proximity of the cofactor and substrate, facilitating the phosphorylation of CDP-ME (Fig. 6) (49, 50).

The binding sites for both CDP-ME and ATP show remarkable conservation in terms of their residues and overall shape. Notably, the cytosine moiety of CDP-ME is situated between two highly conserved aromatic residues with a specific affinity for cytosine, meaning that molecules with adenosine or uridine substituting the cytosine moieties cannot act as substrates. Additionally, the hydroxy group of erythritol forms essential hydrogen bonds with conserved lysine and aspartate side chains, believed to play a crucial role in deprotonation and phosphorylation during catalysis (49, 50).

The ATP binding site is situated in the N-terminal domain, with the adenine base located within an aliphatic crevice, the adenine is maintained into a less favorable syn-conformation through a network of hydrogen bonds. Both the ribose and phosphate components of ATP are exposed to the solvent. The ribose is surrounded by a network of hydrogen bonds, while the phosphate is encircled by a phosphate-binding loop rich in glycine (P-loop). It is worth noting that this P-loop is one of the three highly conserved motifs within the GHMP superfamily, with the other two motifs playing significant roles in shaping the catalytic cavity (Fig. 4) (49–52).

In general, IspE exhibits catalytic mechanism similar to other GHMP family members, the catalysis of IspE involves substrate and cofactor binding in their respective domain's active sites, enabling direct phosphate transfer. Although binding kinetics remain less explored, it likely follows an ordered-sequential process, as seen in similar kinases, where the substrate CDP-ME binds first and the phosphorylated product exits first (53).

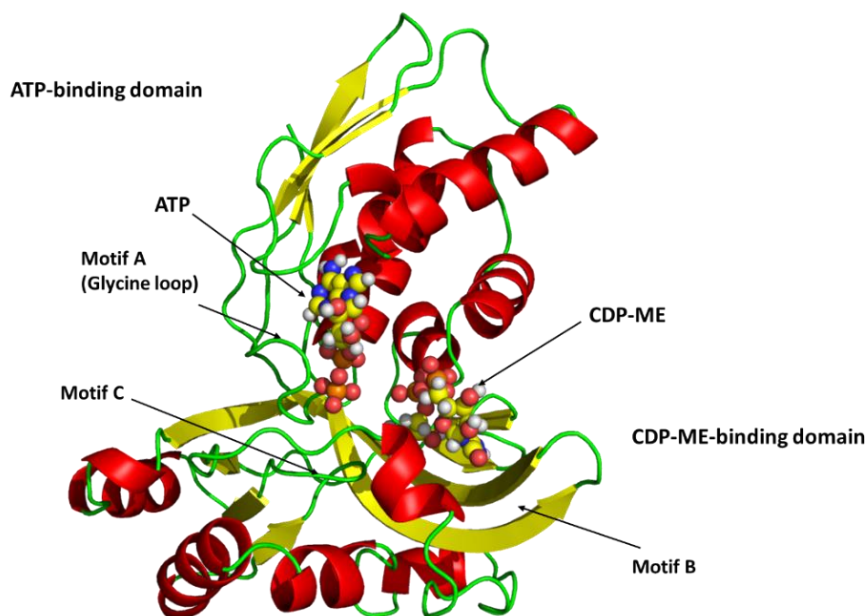


Figure 4 : *EclspE* kinase monomer structure. Highlighting the GHMP conserved motifs.

1.3.2 Targeting 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase

IspE provides two distinct pockets as potential targets in our search for inhibitors: ATP binding site in addition to the substrate binding site, along with nearby sub-pockets. In the context of targeting the ATP binding site, Tang et al. described early ATP competitive inhibitors, they utilized a library of existing GHMP kinases inhibitors. Subsequent studies have adopted similar approaches to identify non-substrate-like inhibitors for IspE. This involved examining existing GHMP kinases inhibitors and employing structure-based drug design coupled with virtual high-throughput screening, utilizing available IspE structures. These efforts led to the discovery of two promising compounds featuring benzene sulfonamide and acetamide moieties (54, 55) (Fig. 5). Over the years, our group has dedicated efforts to investigate both of these pockets. Recently, we conducted a virtual screening on the ATP binding pocket, followed by a Structure-Activity Relationship (SAR) study that resulted in hits with improved pharmacodynamic properties (56). Unfortunately, endeavors thus far haven't yielded highly selective IspE inhibitors specifically targeting the kinase pocket.

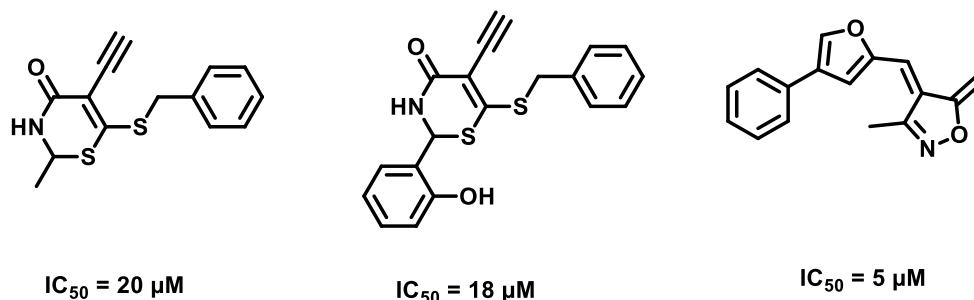


Figure 5: Chemical structures of IspE inhibitors with dual GHMP kinase and IspE activity. Tested against *E. coli* IspE.

IspE inhibitors targeting the substrate pocket appear more promising for further development. In 2007, Hirsch et al. identified compounds with low micromolar IC_{50} values that target this substrate pocket. The Hirsch group played a pioneering role in this area. Their designed inhibitors specifically bound to the cytidine binding site in addition to a deep hydrophobic pocket near phosphate binding site. Importantly, these inhibitors possessed drug-like properties and demonstrated improved activities. Structure-Activity Relationship (SAR) studies revealed the usefulness of this hydrophobic cavity within the CDP-ME binding site, unoccupied by CDP-ME itself. Inhibitor showing highest potency from this class featured a tetrahydrothiophene ring as a ribose substitute, along with a sulfonamide as phosphate substituent attached to a hydrocarbon group to occupy the deep hydrophobic pocket. A co-crystal structure of *Aquifex aeolicus* IspE with one of these compounds confirmed their binding in the substrate-binding pocket (57–59). Subsequently, the group successfully employed bioconjugation to polymer peptide to enhance the solubility and potency of these compounds (Fig. 6) (60).

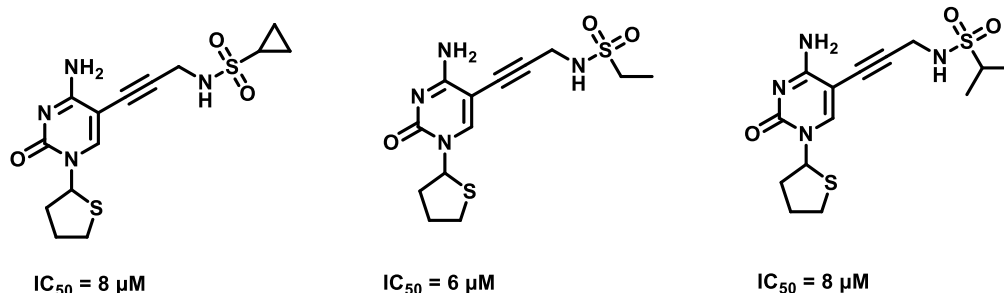


Figure 6: structures of the most potent substrate competitive inhibitors reported by Hirsch et.al (59). IC_{50} measured against *E. coli*.

1.3.3 Co-crystallizing identified LspE kinase inhibitors is a challenging task

A substantial limitation in advancing medicinal chemistry projects aimed at targeting LspE is the absence of a co-crystal structures with pathogenic microorganisms. This is in contrast to the available co-crystal structure with the non-pathogenic bacterium *A. aeolicus*. This can be attributed, at least in part, to the structural conformation of the protein in the asymmetric unit. While experimental findings confirm that LspE kinase predominantly occurs as a monomer in solution (49, 52, 61), during crystallization, there is variability in the number of protein units found in the asymmetric unit. For instance, some LspE structures, such as those from *M. tuberculosis*, crystallize with a single protein unit in the asymmetric unit (52), while others, like two *E.coli* LspE structures (PDB: 1OJ4, 2WW4), display two protein units in the asymmetric unit. Importantly, for 2WW4 the crystal arrangement lacks symmetry, where a loop from each unit obstructs the active site, specifically the CDP-ME binding site of the other (Fig. 7). Subsequent investigations suggest that this crystalline arrangement is an artifact resulting during the crystallization process instead of it being a representation of natural dimerization (62). This common crystalline arrangement, observed in other enzyme homologues as well, presents a considerable challenge for drug design endeavors aiming to co-crystallize or soak LspE inhibitors. Especially those designed to target the substrate binding site, as this pocket becomes inaccessible in this crystalline form, making it difficult to elucidate their binding mode.

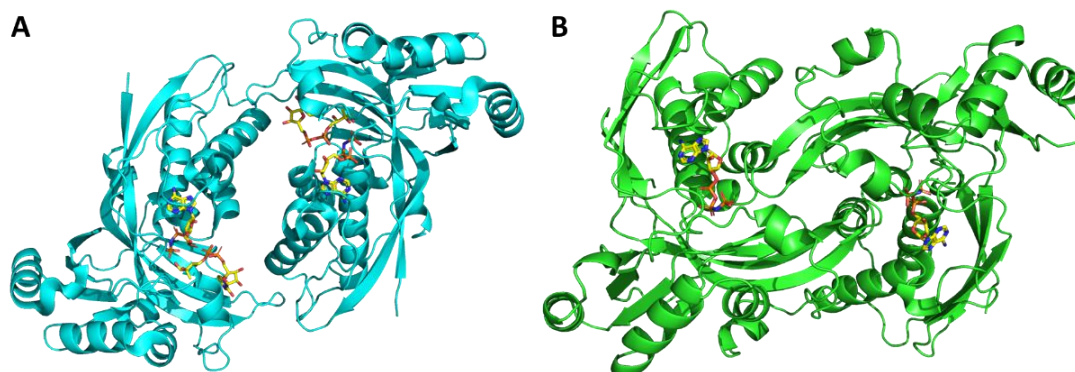


Figure 7: *E.coli* LspE crystal structure different conformations. A; PDB:1oj4, symmetry. B; (PDB: 2ww4) asymmetry.

1.4 Aim of the thesis

The methylerythritol phosphate (MEP) pathway enzymes are promising targets for anti-microbial drugs. Pathogenic microorganisms heavily depend on this pathway to produce vital biological precursors crucial for their cellular processes. Importantly, this pathway is missing in humans, providing the needed selectivity for treating human infections. The main goal of this study is to investigate MEP pathway enzymes as targets for drug design using biochemical and structural biology approaches.

The first three chapters focus on the first enzyme of the MEP pathway, DXPS, which also serves as the rate-limiting step, determining the overall rate of reactions in the pathway. Our aim is to optimize constructs of DXPS to obtain crystal structures from pathogenic organisms, specifically *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* and subsequently acquire co-crystal structures with inhibitors to understand their binding mode, supporting our drug design and discovery projects.

In addition to DXPS we also work with another essential enzyme from the MEP pathway, the IspE kinase. This enzyme catalyzes the fourth step of isoprenoid biosynthesis. Our objective is to investigate the kinetics of IspE from the pathogenic bacteria: *Escherichia coli*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii*. We also concentrate on obtaining and optimizing crystallization conditions and to acquire co-crystal structures with identified hit compounds, aiming to analyze the binding site interactions of IspE in these three homologues, along with their respective inhibitors.

1.5 References

1. World Health Organization (WHO) [online] <https://www.who.int> (Accessed August 8, 2023)
2. Pulingam, T., Parumasivam, T., Gazzali, A. M., Sulaiman, A. M., Chee, J. Y., Lakshmanan, M., Chin, C. F., and Sudesh, K. (2022) Antimicrobial resistance: Prevalence, economic burden, mechanisms of resistance and strategies to overcome. *Eur. J. Pharm. Sci.* **170**, 106103
3. Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993) Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295**, 517–524
4. Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S., and Sahm, H. (1996) Glyceraldehyde 3-Phosphate and Pyruvate as Precursors of Isoprenic Units in an Alternative Non-mevalonate Pathway for Terpenoid Biosynthesis. *J. Am. Chem. Soc.* **118**, 2564–2566
5. Illarionova, V., Kaiser, J., Ostrozhenkova, E., Bacher, A., Fischer, M., Eisenreich, W., and Rohdich, F. (2006) Nonmevalonate Terpene Biosynthesis Enzymes as Anti-infective Drug Targets: Substrate Synthesis and High-Throughput Screening Methods. *J. Org. Chem.* **71**, 8824–8834
6. Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000) Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13172–13177
7. Holstein, S. A., and Hohl, R. J. (2004) Isoprenoids: remarkable diversity of form and function. *Lipids.* **39**, 293–309
8. Sacchettini, J. C., and Poulter, C. D. (1997) Creating Isoprenoid Diversity. *Science.* **277**, 1788–1789
9. Wang, K. C., and Ohnuma, S. (2000) Isoprenyl diphosphate synthases. *Biochim. Biophys. Acta.* **1529**, 33–48
10. Rohmer, M. (2003) Mevalonate-independent methylerythritol phosphate pathway for isoprenoid biosynthesis. Elucidation and distribution. *Pure Appl. Chem.* **75**, 375–388
11. Rohmer, M., and Rohmer, M. (1999) The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants†. *Nat. Prod. Rep.* **16**, 565–574
12. Rodriguez-Concepcion, M. (2004) The MEP Pathway: A New Target for the Development of Herbicides, Antibiotics and Antimalarial Drugs. *Curr. Pharm. Des.* **10**, 2391–2400
13. Zeng, L., and Dehesh, K. (2021) The eukaryotic MEP-pathway genes are evolutionarily conserved and originated from Chlamydia and cyanobacteria. *BMC Genomics.* **22**, 137
14. Wang, X., and Dowd, C. S. (2018) The Methylerythritol Phosphate Pathway: Promising Drug Targets in the Fight against Tuberculosis. *ACS Infect. Dis.* **4**, 278–290
15. Baugh, L., Phan, I., Begley, D. W., Clifton, M. C., Armour, B., Dranow, D. M., Taylor, B. M., Muruthi, M. M., Abendroth, J., Fairman, J. W., Fox, D., Dieterich, S. H., Staker, B. L., Gardberg, A. S., Choi, R., Hewitt, S. N., Napuli, A. J., Myers, J., Barrett, L. K., Zhang, Y., Ferrell, M., Mundt, E., Thompkins, K., Tran, N., Lyons-Abbott, S., Abramov, A., Sekar, A., Serbzhinskiy, D., Lorimer, D., Buchko, G. W.,

- Stacy, R., Stewart, L. J., Edwards, T. E., Van Voorhis, W. C., and Myler, P. J. (2015) Increasing the structural coverage of tuberculosis drug targets. *Tuberculosis*. **95**, 142–148
16. Lichtenthaler, H. K. (2000) Non-mevalonate isoprenoid biosynthesis: enzymes, genes and inhibitors. *Biochem. Soc. Trans.* **28**, 785–789
 17. Wright, L. P., Rohwer, J. M., Ghirardo, A., Hammerbacher, A., Ortiz-Alcaide, M., Raguschke, B., Schnitzler, J.-P., Gershenzon, J., and Phillips, M. A. (2014) Deoxyxylulose 5-Phosphate Synthase Controls Flux through the Methylerythritol 4-Phosphate Pathway in Arabidopsis1[C][W][OPEN]. *Plant Physiol.* **165**, 1488–1504
 18. Moreno-Sánchez, R., Saavedra, E., Rodríguez-Enríquez, S., and Olín-Sandoval, V. (2008) Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways. *J. Biomed. Biotechnol.* **2008**, 597913
 19. Volke, D. C., Rohwer, J., Fischer, R., and Jennewein, S. (2019) Investigation of the methylerythritol 4-phosphate pathway for microbial terpenoid production through metabolic control analysis. *Microb. Cell Factories*. **18**, 192
 20. Lell, B., Ruangweeraayut, R., Wiesner, J., Missinou, M. A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H., and Kremsner, P. G. (2003) Fosmidomycin, a Novel Chemotherapeutic Agent for Malaria. *Antimicrob. Agents Chemother.* **47**, 735–738
 21. Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H. K., Soldati, D., and Beck, E. (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science*. **285**, 1573–1576
 22. Oyakhirome, S., Issifou, S., Pongratz, P., Barondi, F., Ramharter, M., Kun, J. F., Missinou, M. A., Lell, B., and Kremsner, P. G. (2007) Randomized Controlled Trial of Fosmidomycin-Clindamycin versus Sulfadoxine-Pyrimethamine in the Treatment of Plasmodium falciparum Malaria. *Antimicrob. Agents Chemother.* **51**, 1869–1871
 23. Jomaa Pharma GmbH (2015) *A Phase IIa Proof of Concept Study to Explore the Efficacy, Tolerability and Safety of Fosmidomycin Sodium When Administered With Piperaquine Tetraphosphate to Adults and Older Children With Acute Uncomplicated Plasmodium Falciparum Malaria*, clinicaltrials.gov
 24. Sprenger, G. A., Schörken, U., Wiegert, T., Grolle, S., de Graaf, A. A., Taylor, S. V., Begley, T. P., Bringer-Meyer, S., and Sahm, H. (1997) Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12857–12862
 25. Xiang, S., Usunow, G., Lange, G., Busch, M., and Tong, L. (2007) Crystal Structure of 1-Deoxy-d-xylulose 5-Phosphate Synthase, a Crucial Enzyme for Isoprenoids Biosynthesis. *J. Biol. Chem.* **282**, 2676–2682
 26. Hamid, R., Adam, S., Lacour, A., Monjas, L., Köhnke, J., and Hirsch, A. K. H. (2023) 1-deoxy-D-xylulose-5-phosphate synthase from Pseudomonas aeruginosa and Klebsiella pneumoniae reveals conformational changes upon cofactor binding. *J. Biol. Chem.* **299**, 105152
 27. Vogel, C., and Pleiss, J. (2014) The modular structure of ThDP-dependent enzymes. *Proteins*. **82**, 2523–2537

28. Brammer Basta, L. A., Patel, H., Kakalis, L., Jordan, F., and Freil Meyers, C. L. (2014) Defining critical residues for substrate binding to 1-deoxy-d-xylulose 5-phosphate synthase – active site substitutions stabilize the predecarboxylation intermediate C2 α -lactylthiamin diphosphate. *FEBS J.* **281**, 2820–2837
29. Querol, J., Rodríguez-Concepción, M., Boronat, A., and Imperial, S. (2001) Essential role of residue H49 for activity of Escherichia coli 1-deoxy-D-xylulose 5-phosphate synthase, the enzyme catalyzing the first step of the 2-C-methyl-D-erythritol 4-phosphate pathway for isoprenoid Synthesis. *Biochem. Biophys. Res. Commun.* **289**, 155–160
30. Brammer, L. A., Smith, J. M., Wade, H., and Meyers, C. F. (2011) 1-Deoxy-d-xylulose 5-Phosphate Synthase Catalyzes a Novel Random Sequential Mechanism. *J. Biol. Chem.* **286**, 36522–36531
31. Matsue, Y., Mizuno, H., Tomita, T., Asami, T., Nishiyama, M., and Kuzuyama, T. (2010) The herbicide ketoclofomazone inhibits 1-deoxy-D-xylulose 5-phosphate synthase in the 2-C-methyl-D-erythritol 4-phosphate pathway and shows antibacterial activity against Haemophilus influenzae. *J. Antibiot. (Tokyo)*. **63**, 583–588
32. Patel, H., Nemeria, N. S., Brammer, L. A., Freil Meyers, C. L., and Jordan, F. (2012) Observation of Thiamin-Bound Intermediates and Microscopic Rate Constants for Their Interconversion on 1-Deoxy- D -xylulose 5-Phosphate Synthase: 600-Fold Rate Acceleration of Pyruvate Decarboxylation by D -Glyceraldehyde-3-phosphate. *J. Am. Chem. Soc.* **134**, 18374–18379
33. DeColli, A. A., Nemeria, N. S., Majumdar, A., Gerfen, G. J., Jordan, F., and Freil Meyers, C. L. (2018) Oxidative decarboxylation of pyruvate by 1-deoxy-d-xylulose 5-phosphate synthase, a central metabolic enzyme in bacteria. *J. Biol. Chem.* **293**, 10857–10869
34. Handa, S., Dempsey, D. R., Ramamoorthy, D., Cook, N., Guida, W. C., Spradling, T. J., White, J. K., Woodcock, H. L., and Merkler, D. J. (2018) Mechanistic Studies of 1-Deoxy-D-Xylulose-5-Phosphate Synthase from Deinococcus radiodurans. *Biochem. Mol. Biol. J.* **4**, 2
35. DeColli, A. A., Zhang, X., Heflin, K. L., Jordan, F., and Freil Meyers, C. L. (2019) Active Site Histidines Link Conformational Dynamics with Catalysis on Anti-Infective Target 1-Deoxy- D -xylulose 5-Phosphate Synthase. *Biochemistry*. **58**, 4970–4982
36. Chen, P. Y.-T., DeColli, A. A., Meyers, C. L. F., and Drennan, C. L. (2019) X-ray crystallography–based structural elucidation of enzyme-bound intermediates along the 1-deoxy-d-xylulose 5-phosphate synthase reaction coordinate. *J. Biol. Chem.* **294**, 12405–12414
37. Hayashi, D., Kato, N., Kuzuyama, T., Sato, Y., and Ohkanda, J. (2013) Antimicrobial N-(2-chlorobenzyl)-substituted hydroxamate is an inhibitor of 1-deoxy-D-xylulose 5-phosphate synthase. *Chem. Commun.* **49**, 5535–5537
38. Masini, T., and Hirsch, A. K. H. (2014) Development of Inhibitors of the 2C-Methyl-d-erythritol 4-Phosphate (MEP) Pathway Enzymes as Potential Anti-Infective Agents. *J. Med. Chem.* **57**, 9740–9763
39. Zhu, D., Johannsen, S., Masini, T., Simonin, C., Haupenthal, J., Illarionov, B., Andreas, A., Awale, M., Gierse, R. M., Laan, T. van der, Vlag, R. van der, Nasti, R., Poizat, M., Buhler, E., Reiling, N., Müller, R., Fischer, M., Reymond, J.-L., and Hirsch, A. K. H. (2022) Discovery of novel drug-like antitubercular hits targeting the MEP pathway enzyme DXPS by strategic application of ligand-based virtual screening. *Chem. Sci.* **13**, 10686–10698

40. Nemeria, N. S., Shome, B., DeColli, A. A., Heflin, K., Begley, T. P., Meyers, C. F., and Jordan, F. (2016) Competence of Thiamin Diphosphate-Dependent Enzymes with 2'-Methoxythiamin Diphosphate Derived from Bacimethrin, a Naturally Occurring Thiamin Anti-vitamin. *Biochemistry*. **55**, 1135–1148
41. Smith, J. M., Vierling, R. J., and Meyers, C. F. (2012) Selective inhibition of E. coli 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *Med Chem Commun*. **3**, 65–67
42. Smith, J. M., Warrington, N. V., Vierling, R. J., Kuhn, M. L., Anderson, W. F., Koppisch, A. T., and Freil Meyers, C. L. (2014) Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *J. Antibiot. (Tokyo)*. **67**, 77–83
43. Barteel, D., Sanders, S., Phillips, P. D., Harrison, M. J., Koppisch, A. T., and Meyers, C. L. F. (2019) Enamide Prodrugs of Acetyl Phosphonate Deoxy-d-xylulose-5-phosphate Synthase Inhibitors as Potent Antibacterial Agents. *ACS Infect. Dis.* 10.1021/acsinfecdis.8b00307
44. Gierse, R. M., Reddem, E. R., Alhayek, A., Baitinger, D., Hamid, Z., Jakobi, H., Laber, B., Lange, G., Hirsch, A. K. H., and Groves, M. R. (2021) Identification of a 1-deoxy-D-xylulose-5-phosphate synthase (DXS) mutant with improved crystallographic properties. *Biochem. Biophys. Res. Commun.* **539**, 42–47
45. Gierse, R. M., Oerlemans, R., Reddem, E. R., Gawriljuk, V. O., Alhayek, A., Baitinger, D., Jakobi, H., Laber, B., Lange, G., Hirsch, A. K. H., and Groves, M. R. (2022) First crystal structures of 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) from Mycobacterium tuberculosis indicate a distinct mechanism of intermediate stabilization. *Sci. Rep.* **12**, 7221
46. Gawriljuk, V. O., Oerlemans, R., Gierse, R. M., Jotwani, R., Hirsch, A. K. H., and Groves, M. R. (2023) Structure of Mycobacterium tuberculosis 1-Deoxy-D-Xylulose 5-Phosphate Synthase in Complex with Butylacetylphosphonate. *Crystals*. **13**, 737
47. Lüttgen, H., Rohdich, F., Herz, S., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., Fellermeier, M., Sagner, S., Zenk, M. H., Bacher, A., and Eisenreich, W. (2000) Biosynthesis of terpenoids: YchB protein of Escherichia coli phosphorylates the 2-hydroxy group of 4-diphosphocytidyl-2C-methyl-d-erythritol. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1062–1067
48. Kuzuyama, T., Takagi, M., Kaneda, K., Watanabe, H., Dairi, T., and Seto, H. (2000) Studies on the nonmevalonate pathway: conversion of 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol to its 2-phospho derivative by 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase. *Tetrahedron Lett.* **41**, 2925–2928
49. Wada, T., Kuzuyama, T., Satoh, S., Kuramitsu, S., Yokoyama, S., Unzai, S., Tame, J. R. H., and Park, S.-Y. (2003) Crystal Structure of 4-(Cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase, an Enzyme in the Non-mevalonate Pathway of Isoprenoid Synthesis*. *J. Biol. Chem.* **278**, 30022–30027
50. Miallau, L., Alphey, M. S., Kemp, L. E., Leonard, G. A., McSweeney, S. M., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F., and Hunter, W. N. (2003) Biosynthesis of isoprenoids: Crystal structure of 4-diphosphocytidyl-2C-methyl-d-erythritol kinase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9173–9178
51. Sgraja, T., Alphey, M. S., Ghilagaber, S., Marquez, R., Robertson, M. N., Hemmings, J. L., Lauw, S., Rohdich, F., Bacher, A., Eisenreich, W., Illarionova, V., and Hunter, W. N. (2008) Characterization

- of *Aquifex aeolicus* 4-diphosphocytidyl-2C-methyl-d-erythritol kinase – ligand recognition in a template for antimicrobial drug discovery. *FEBS J.* **275**, 2779–2794
52. Shan, S., Chen, X., Liu, T., Zhao, H., Rao, Z., and Lou, Z. (2011) Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*. *FASEB J.* **25**, 1577–1584
 53. Fu, Z., Wang, M., Potter, D., Mizioroko, H. M., and Kim, J.-J. P. (2002) The Structure of a Binary Complex between a Mammalian Mevalonate Kinase and ATP: INSIGHTS INTO THE REACTION MECHANISM AND HUMAN INHERITED DISEASE*. *J. Biol. Chem.* **277**, 18134–18142
 54. Tang, M., Odejinmi, S., Allette, Y., Vankayalapati, H., and Lai, K. (2011) Identification of Novel Small Molecule Inhibitors of 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) kinase of Gram-negative bacteria. *Bioorg. Med. Chem.* **19**, 5886–5895
 55. Tidten-Luksch, N., Grimaldi, R., Torrie, L. S., Frearson, J. A., Hunter, W. N., and Brenk, R. (2012) IspE Inhibitors Identified by a Combination of In Silico and In Vitro High-Throughput Screening. *PLoS ONE.* **7**, e35792
 56. Ropponen, H.-K., Diamanti, E., Johannsen, S., Illarionov, B., Hamid, R., Jaki, M., Sass, P., Fischer, M., Hauptenthal, J., and Hirsch, A. K. H. Exploring the translational gap of a novel class of *Escherichia coli* IspE inhibitors. *ChemMedChem.* **n/a**, e202300346
 57. Hirsch, A. K. H., Lauw, S., Gersbach, P., Schweizer, W. B., Rohdich, F., Eisenreich, W., Bacher, A., and Diederich, F. (2007) Nonphosphate Inhibitors of IspE Protein, a Kinase in the Non-Mevalonate Pathway for Isoprenoid Biosynthesis and a Potential Target for Antimalarial Therapy. *ChemMedChem.* **2**, 806–810
 58. Crane, C. M., Hirsch, A. K. H., Alphey, M. S., Sgraja, T., Lauw, S., Illarionova, V., Rohdich, F., Eisenreich, W., Hunter, W. N., Bacher, A., and Diederich, F. (2008) Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis. *ChemMedChem.* **3**, 91–101
 59. H. Hirsch, A. K., S. Alphey, M., Lauw, S., Seet, M., Barandun, L., Eisenreich, W., Rohdich, F., N. Hunter, W., Bacher, A., and Diederich, F. (2008) Inhibitors of the kinase IspE: structure–activity relationships and co-crystal structure analysis. *Org. Biomol. Chem.* **6**, 2719–2730
 60. Hirsch, A. K. H., Diederich, F., Antonietti, M., and Börner, H. G. (2009) Bioconjugates to specifically render inhibitors water-soluble. *Soft Matter.* **6**, 88–91
 61. Frank, A., and Groll, M. (2017) The Methylerythritol Phosphate Pathway to Isoprenoids. *Chem. Rev.* **117**, 5675–5703
 62. Kalinowska-Tłuścik, J., Miallau, L., Gabrielsen, M., Leonard, G. A., McSweeney, S. M., and Hunter, W. N. (2010) A triclinic crystal form of *Escherichia coli* 4-diphosphocytidyl-2 C -methyl- D -erythritol kinase and reassessment of the quaternary structure. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **66**, 237–241

Chapter 2

Results and discussion

2 RESULTS AND DISCUSSIONS

The results and discussions chapter will be divided into the four sections listed below. The first two and the last parts have been published or submitted to a journal, and they will be presented in accordance with the formatting of their respective journals. Part three, which is part of an ongoing work, will be presented as a short report.

2.1 1-Deoxy-D-xylulose-5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding

2.1.1 Supplementary information

2.2 Design of thiamine analogues for inhibition of thiamine diphosphate (ThDP)-dependent enzymes: Systematic investigation through Scaffold-Hopping and C2-Functionalisation

2.2.1 Supplementary information

2.3 Allosteric Inhibition of DXPS: A New Approach in Drug Discovery

2.4 IspE Kinase Homologues: Characterization and crystal Structures enable structure-based design of novel inhibitors

2.4.1 Supplementary information

Chapter 2.1

1-Deoxy-D-xylulose-5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding



1-deoxy-D-xylulose-5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding

Received for publication, June 11, 2023, and in revised form, August 1, 2023. Published, Papers in Press, August 9, 2023.

<https://doi.org/10.1016/j.jbc.2023.105152>

Rawia Hamid^{1,2}, Sebastian Adam¹ , Antoine Lacour¹ , Leticia Monjas³, Jesko Köhnke^{4,5} , and Anna K. H. Hirsch^{1,2,*} 

From the ¹Department of Drug Design and Optimization, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Saarbrücken, Germany; ²Department of Pharmacy, Saarland University, Saarbrücken, Germany; ³Stratingh Institute for Chemistry, University of Groningen, Groningen, The Netherlands; ⁴Institute of Food Chemistry, Leibniz University Hannover, Hannover, Germany; ⁵School of Chemistry, University of Glasgow, Glasgow, UK

Reviewed by members of the JBC Editorial Board. Edited by Joseph Jez

The ESKAPE bacteria are the six highly virulent and antibiotic-resistant pathogens that require the most urgent attention for the development of novel antibiotics. Detailed knowledge of target proteins specific to bacteria is essential to develop novel treatment options. The methylerythritol-phosphate (MEP) pathway, which is absent in humans, represents a potentially valuable target for the development of novel antibiotics. Within the MEP pathway, the enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXPS) catalyzes a crucial, rate-limiting first step and a branch point in the biosynthesis of the vitamins B1 and B6. We report the high-resolution crystal structures of DXPS from the important ESKAPE pathogens *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* in both the co-factor-bound and the apo forms. We demonstrate that the absence of the cofactor thiamine diphosphate results in conformational changes that lead to disordered loops close to the active site that might be important for the design of potent DXPS inhibitors. Collectively, our results provide important structural details that aid in the assessment of DXPS as a potential target in the ongoing efforts to combat antibiotic resistance.

The widespread use of antibiotics over the past five decades has caused bacteria to develop resistance mechanisms to evade the detrimental effect of these agents (1, 2). This effect is especially critical for infections caused by ESKAPE pathogens, which are difficult to treat and result in increasing fatality rates. *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* are both Gram-negative pathogens with the ability to integrate exogenous DNA to obtain antibiotic resistance (horizontal gene transfer) (3). Both organisms can adapt to the environment of human airways (4) and are major causes of opportunistic infections, predominantly, pneumonia and sepsis in hospitalized patients and patients with cystic fibrosis (5, 6). Measures such as the antibiotic stewardship program to improve the way antibiotics are prescribed, and the synthesis

of novel chemical entities with new modes of action need to be implemented to tackle the antimicrobial-resistance crisis.

The methylerythritol-phosphate (MEP) pathway has attracted attention as it represents a rich source of potentially attractive anti-infective drug targets. The MEP pathway, also called the mevalonate-independent pathway, is considered to be the main pathway for the synthesis of isoprenoid building blocks in most bacteria, plants, and protozoa (7–9). Its absence in humans makes the enzymes of this pathway, particularly interesting targets for anti-infective drug discovery. The complete pathway consists of seven enzymes, each of which could be targeted individually. 1-Deoxy-D-xylulose-5-phosphate synthase (DXPS) catalyzes the first rate-limiting step of the MEP pathway. The enzyme is responsible for the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) by condensation of pyruvate and glyceraldehyde 3-phosphate (D-GAP) in the presence of thiamine diphosphate (ThDP), Figure 1. DXP, in and of itself, is an important metabolite in not only the biosynthesis of isoprenoid precursors but also in vitamin B1 (thiamine) and vitamin B6 (pyridoxine) biosynthesis (10). DXPS is a promising anti-infective drug target (11–13).

The first DXPS structures of *Deinococcus radiodurans* and *Escherichia coli* were solved via X-ray crystallography and published in 2007 (14). Subsequent studies on these structures have shown that DXPS undergoes a number of conformational changes during catalysis which takes place near the substrate-entry site (15–17). The ThDP-bound enzyme, in the presence of pyruvate, forms a pre-decarboxylation intermediate, lactyl thiamine diphosphate (LThDP) (15, 18). The enzyme adopts a closed conformation when bound to LThDP and exposure to D-GAP or DXP triggers a conformational change to an open conformation, which seemingly initiates LThDP decarboxylation (17, 19). Chen *et al.* coined the name spoon-fork motif to describe the shape of the region that shows conformational changes during catalysis. They have shown that binding of LThDP results in a closed conformation of the spoon motif that opens when the subsequent intermediate enamine is formed (17). These results highlight the

* For correspondence: Anna K. H. Hirsch, anna.hirsch@helmholtz-hips.de.

DXPS conformational changes upon cofactor binding

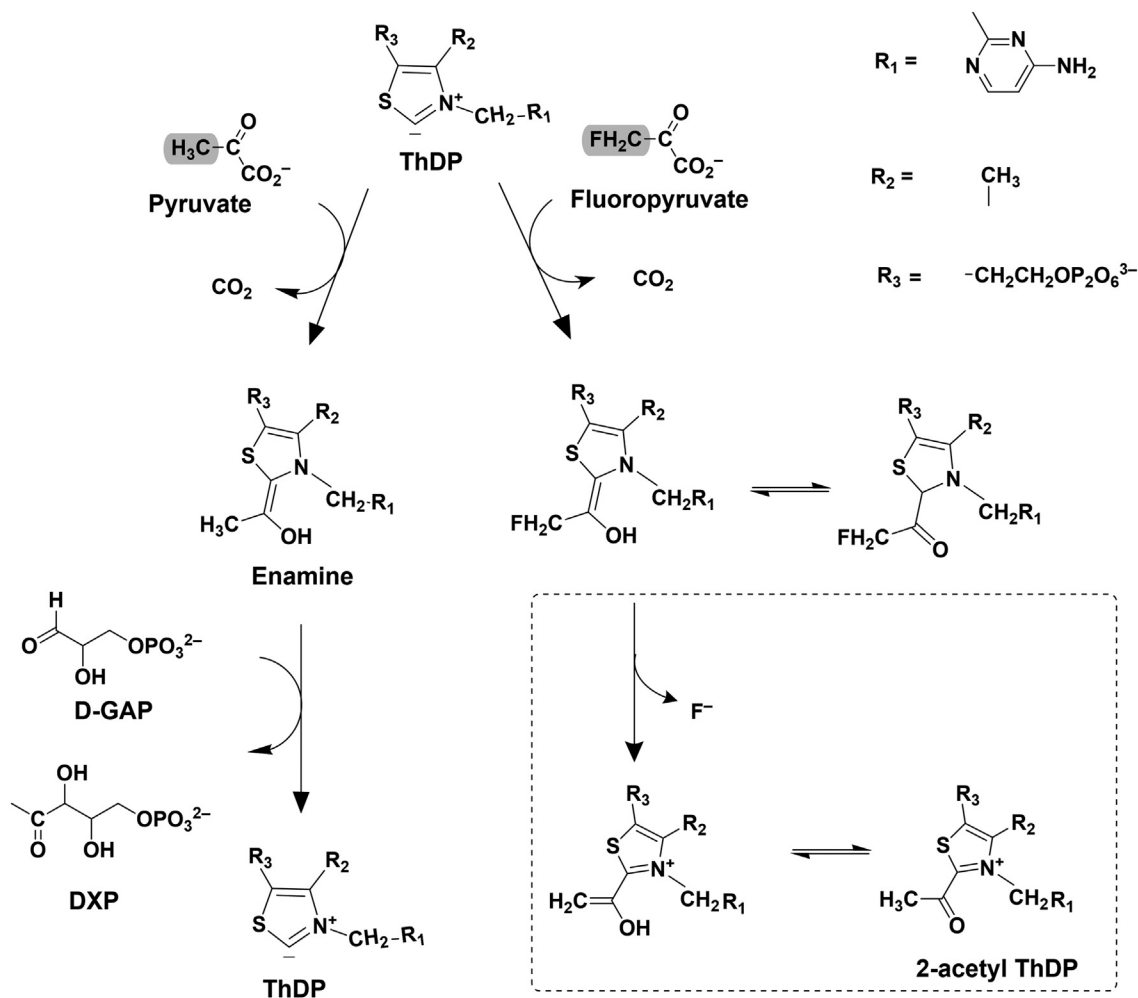


Figure 1. DXPS reaction mechanism with pyruvate and fluoropyruvate. Suggested formation of the fluoropyruvate adduct (2-acetyl ThDP) catalyzed by *pa*DXPS in combination with ThDP in comparison to the reported reaction with the natural substrate pyruvate. The point of difference between fluoropyruvate and pyruvate is highlighted in gray.

important role of conformational changes in DXPS catalysis. Further investigation is needed to expand our knowledge about the dynamic catalysis and structural differences of this unique enzyme.

Given that the binding site of the cofactor ThDP is the most promising site for drugs targeting ThDP-dependent enzymes including DXPS (16, 20, 21), it is also vital to understand the precise location of ThDP and the general three-dimensional shape of the pocket. It is also considered the main druggable pocket in DXPS enzymes (22). Thus far, few experimentally determined structures of DXPS have been reported (23), but due to the generally conserved active sites, the available crystal structures of DXPS are used as models for structure-based drug design (16, 20, 21). However, this has proven to be very limiting, as three-dimensional active sites can typically not be deduced from sequence homology alone. To improve our understanding of DXPS active-site architecture and to provide a better framework for drug-discovery efforts addressing this promising target, we determined the crystal structures of both the *apo* and the cofactor-bound forms from the ESKAPE pathogens *P. aeruginosa* and *K. pneumoniae*. We demonstrate

that the binding of ThDP leads to conformational changes involving a loop in domain I, which alters the ThDP binding site and contributes to the binding of the ThDP diphosphate. Together with the structure of *P. aeruginosa* DXPS bound to the less reactive 2-acetylThDP, we provide a firm basis for future drug design and -discovery efforts targeting this important enzyme from the MEP pathway (2).

Results and discussion

Expression, purification, and kinetic characterization of DXPS

Our group recently reported a truncated *D. radiodurans* DXPS (*dr*DXPS) construct in which a flexible loop has been modified to produce a protein that crystallizes readily and yields high-quality crystals that diffract to a resolution better than 2.0 Å (24). This loop typically has very low evolutionary conservation and neither shapes the active site nor is it involved in the formation of a secondary structure (24). Our construct served as a blueprint for the alteration in DXPS enzymes from different organisms (23). After multiple sequence alignments of published sequences, we identified the

loop to be corresponding to residues 206 to 245 in *paDXPS* and residues 197 to 240 in *kpDXPS* (Fig. S1) and replaced them with a six glycine unit. We then heterologously expressed the enzymes in *E. coli*, purified and analyzed them as detailed in the Experimental procedures section. Next, we analyzed the activity of new homolog constructs, evaluated, and compared the enzyme kinetics for both versions of the enzyme *i.e.* truncated and full-length, of the enzyme. The results summarized in Table 1 demonstrate comparable activities of the native and truncated *paDXPS*, both versions of the enzyme show similar affinities for the substrate and co-factor. Although native *kpDXPS* was not tested, the truncated version is showing comparable kinetics for both substrates. No kinetics values for ThDP were reported for other homologs, which makes our data impossible to compare to literature values. However, *kpDXPS* shows a slightly higher affinity for the substrates and elevated catalytic activity in comparison to *paDXPS*, which may be due to the different shape of the binding site. The kinetic parameters remain within a similar range for *paDXPS*s, and the preserved enzymatic activity of *kpDXPS* suggests that the truncation does not affect any catalytically significant residues. These results allowed us to use these optimized and more stable constructs for biological testing as well as structure determination. After confirming the identity and activity of our purified enzymes, we opted to elucidate the structures of these novel homologs from ESKAPE pathogens.

Structure determination of DXPS from *K. pneumoniae* and *P. aeruginosa*

The optimized constructs of DXPS from *K. pneumoniae* and *P. aeruginosa* were used for crystallization trials. The structures were determined using molecular replacement with the published DXPS structure of *E. coli* (PDB ID: 2o1s) as a search model. The two proteins in the asymmetric unit form the expected biological dimer known from other DXPS enzymes, which was supported by native MS measurements for *paDXPS* (Fig. S2). The low solubility of *kpDXPS* precluded native MS data acquisition, but an analysis of the dimer observed in the crystal structure using the PISA server (25) and analytical gel filtration (Fig. S3) confirmed a stable dimer interface, which is highly similar between all DXPS structures. Although we determined a comparable affinity for ThDP, we can only observe electron density for ThDP in one of the two protomers of the dimer of *kpDXPS* with low occupancy. This may be the result of the longer crystal formation time of 20 days and the

consequent degradation of the cofactor. Intriguingly, the N-terminus of this protomer extends into the cofactor binding site of the dimer partner up to Ser186. Given that this would lead to an inactive enzyme and the absence of this arrangement from the apo structures (see below), we consider this arrangement a crystallographic artifact representing a low-energy state of this particular crystal form. The addition of ThDP was essential to obtain this crystal form, and we thus screened new conditions to obtain an apo structure. Crystals of *apo kpDXPS* were obtained at a considerably higher concentration than the cofactor-bound version and crystals grew more quickly. They belonged to space group C 222₁ and a complete 2.1 Å dataset was collected at ESRF beamline ID23 to 1. The structure was solved using molecular replacement with a refined ThDP-bound *kpDXPS* structure as the search model.

Similar to *kpDXPS*, we chose an optimized construct to obtain a crystal structure of DXPS from *P. aeruginosa* (*paDXPS*). The structure was determined to 2.0 Å with molecular replacement using the published structure of *drDXPS* (PDB ID: 6OUV) as a search model. Subsequently, a refined *paDXPS* structure was used as the search model for the molecular replacement of structures of ThDP-bound *paDXPS* (2.3 Å), a structure with 2-acetylThDP (2.0 Å), and a co-crystal structure with a selected ThDP analog (2.2 Å). In general, *paDXPS* crystallized in orthorhombic crystal form, space group P2₁2₁2₁. The asymmetric unit of all *paDXPS* structures consisted of six protomers, four of which formed dimers with proteins within the asymmetric unit, while the remaining two protomers formed dimers with symmetry mates. In most obtained *paDXPS* structures, approximately the first 30 N-terminal residues are disordered. Native MS data on *paDXPS* confirmed that the enzyme exists as a dimer in solution (Fig. S2), which is consistent with the crystallographic results and suggests that the dimer is the biological unit *in vivo*. All data collection and refinement statistics can be found in Table S2, and the structures were deposited in the PDB.

Structural comparison of *kpDXPS* and *paDXPS*

kpDXPS and *paDXPS* share high sequence homology with previously identified homologs from *E. coli* (*ecDXPS*) and *D. radiodurans* (*drDXPS*) and *Mycobacterium tuberculosis* (*mtDXPS*) (14, 23). *kpDXPS* and *paDXPS* share 62.5% sequence identity and 84.3% sequence similarity. Consequently, the structures are also highly similar with a C_α-RMSD of only 0.716 when comparing both ThDP-bound crystal structures. Each DXPS monomer in all structures we determined consists of three

Table 1
The steady-state kinetics parameters of ThDP, pyruvate, and D-GAP for the wild-type and mutated *paDXPS* and mutated *kpDXPS*

Enzyme	Native <i>paDXPS</i> 0.15 μmol/L	Truncated <i>paDXPS</i> 0.15 μmol/L	Truncated <i>kpDXPS</i> 0.20 μmol/L
Pyruvate	K_m : 184 ± 25 μM K_{cat} : 1.6 ± 0.01 × 10 ³ min ⁻¹	K_m : 122 ± 8.2 μM K_{cat} : 1.2 ± 0.02 × 10 ³ min ⁻¹	K_m : 113 ± 34 μM K_{cat} : 2.1 ± 0.03 × 10 ³ min ⁻¹
D-GAP	K_m : 357 ± 77 μM K_{cat} : 1.8 ± 0.02 × 10 ³ min ⁻¹	K_m : 336 ± 98.0 K_{cat} : 1.5 ± 0.03 × 10 ³ min ⁻¹	K_m : 170 ± 43 μM K_{cat} : 2.4 ± 0.01 × 10 ³ min ⁻¹
ThDP	K_m : 70 ± 9.0 nM K_{cat} : 1.6 ± 0.01 × 10 ³ min ⁻¹	K_m : 96.5 ± 14 nM K_{cat} : 0.94 ± 0.01 × 10 ³ min ⁻¹	K_m : 110 ± 17.0 nM K_{cat} : 2.6 ± 0.03 × 10 ³ min ⁻¹

When measuring K_m , the other substrate and/or cofactor concentrations were kept in excess to ensure the enzyme reaches maximum velocity.

DXPS conformational changes upon cofactor binding

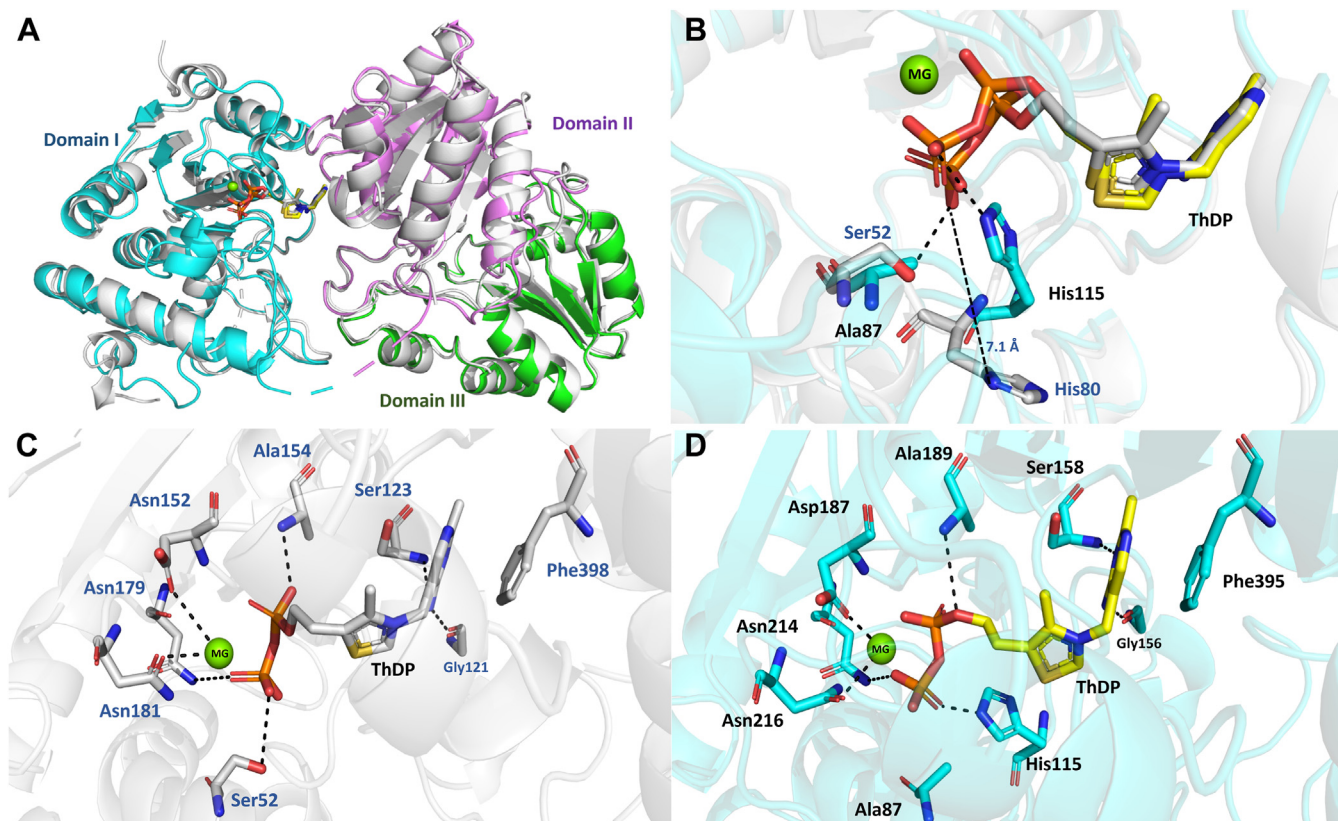


Figure 2. Overview of the overall structures of *pa*DXPS and *kp*DXPS. A, domain organization; domain I (cyan; residues 1–306) contains a five-stranded parallel β -sheet, domain II (violet; residues 317–493) contains a six-stranded parallel β -sheet, domain III (green; residues 496–620) contains a five-stranded β -sheet, in *pa*DXPS, aligned with *kp*DXPS structure in gray. B, comparison of ThDP-bound *kp*DXPS and *pa*DXPS structures highlighting differences in the active site. *pa*DXPS is shown in cyan, overlaid with *kp*DXPS in gray, different residues from *kp*DXPS are labeled in blue. C and D, important interactions with ThDP in *pa*DXPS and *kp*DXPS, respectively. Interactions are shown as black, dashed lines (below 5 Å). All figures of this type were created in PyMol.

domains. The domain organization is shown in Figure 2A and is highly similar to the other homologs. The loop connecting domains I and II (residues 307–315 in *pa*DXPS) is highly flexible and appears to be disordered in most structures. Similar to other reported homologs, the binding site of ThDP is located at the interface between domains I and II with the pyrimidine ring facing domain II and the diphosphate moiety pointing toward domain I. ThDP adopts a chair-like conformation; the pyrimidine ring is perpendicular to the thiazole ring, which is, in turn, perpendicular to the diphosphate. In the *pa*DXPS structure, the pyrimidine ring of ThDP is π - π stacked against Phe395, while N3 of the pyrimidine is forming a backbone hydrogen bond (2.9 Å) with Ser158. Residues His115, Asp187, Ala189, Asn214, Asn216, and Lys286 are coordinating the diphosphate/Mg²⁺ part of ThDP (Fig. 2D). The active site of *kp*DXPS shows an almost identical binding mode of ThDP, and similar residues are involved in binding. The diphosphate is coordinated in an asparagine-rich region, with the contribution of Ser52, which forms an H bond with the diphosphate. This residue is an alanine in the *pa*DXPS structure. Among other structurally characterized DXPS homologs, *pa*DXPS, together with *mt*DXPS, seem to be the only homologs in which this hydrogen bond cannot be formed. In *mt*DXPS, this residue is changed to proline (Fig. S4). Further comparison of the two DXPS structures, reveals a key difference in the active site. His115, in *pa*DXPS, is positioned to form a hydrogen bond with the diphosphate of ThDP (Fig. 2B).

This interaction cannot be formed in *kp*DXPS as the histidine side chain is now 7 Å away from the diphosphate (Fig. 2B). This may be caused by Ser52 preventing His80 from adopting the conformation seen in *pa*DXPS (steric hindrance). The H bonding interaction with His115 specific to *pa*DXPS seems to be compensating for the loss of the interaction with diphosphate caused by the change from serine to alanine in position 87 (Fig. 2D). This residue might be an important distinction for drug design, as the hydroxyl group of serine adds a targetable nucleophilic moiety to the active site. Fig. S4 shows the active site differences between the structures of the published DXPS homologs from *E. coli*, *D. radiodurans*, and *M. tuberculosis* and our structures. Of these, the most noteworthy is the change of Ala189 placed to make direct backbone interaction with the N2 of the pyrimidine in the homolog from the non-pathogenic bacterium *D. radiodurans* to serine (Ser158 in *pa*DXPS and Ser123 in *kp*DXPS, Fig. 2, C and D) as well as the homologs from the pathogenic *E. coli*, where the hydroxyl group of the serine side chain can again be targeted by hydrogen bonding with a ligand.

Structural analysis of the conformational difference in the binding site between ThDP-bound and apo DXPS

The influence of cofactor binding on the shape of the active site was evaluated by obtaining structural information on the apo *pa*DXPS and *kp*DXPS and comparing it to the cofactor-

bound proteins. The loop spanning from Asn216 to Trp250, contributing to the ThDP binding site, exhibits a conformational change between the *holo* and the *apo* structure of *paDXPS* (Fig. 3A). In the cofactor-bound version, residues Asp217, Met218, Ser219, and Ile220 fold around the diphosphate moiety (Fig. 3). The coordination of the co-factor bound Mg^{2+} by the side chain of Asn216 and the main chain of Met218 appear to stabilize the loop. The ThDP-bound structure also shows the participation of the Ser219 and Ile220 side chains, with the latter showing the potential to form an arene-H interaction with the thiazole ring (Fig. 3D). Most other DXPS structures exhibit disorder in this particular region. Furthermore, our study presents the first report of an *apo* DXPS structure, making it challenging to directly compare our results with those of other homologs. In the *kpDXPS* structure regions near the active site seem to undergo various conformational changes. We observe loop rearrangement around the active site: Three loops are partially occluding the active site in the *apo* structure, and appear to get displaced by the bound cofactor, resulting in larger movements averaging around 4 Å (Fig. 4). The loop belonging to the domain I of DXPS (amino acids 91–124) is completely disordered in the *apo* structure of *kpDXPS*, leading to a more open conformation towards the other side of the substrate entrance channel (pale green region in (Fig. 4, A and B). Unfortunately, this part of the protein could

not be resolved sufficiently to allow us to compare it to the structure of *paDXPS*. Furthermore, while the structure obtained for the *apo kpDXPS* has a Rfree value of 20%, indicating a reasonable fit to the experimental data, it is important to state that the Wilson B factor is relatively high at 35 Å². This could be linked to the disordered regions observed in the absence of the cofactor, showcasing the role of ThDP in stabilizing the overall structure of DXPS. The distinct behavior of the loops surrounding the active sites of the two homologs may have important implications for inhibitor design, as a different three-dimensional fold needs to be targeted.

The structure of *paDXPS* with thiamine analog provides evidence for the contribution of the diphosphate moiety in stabilizing a loop close to the active site

Thiamine analogs were previously reported as potent inhibitors of ThDP-dependent enzymes (26, 27). Several thiamine analogs had been synthesized in our group as inhibitors of DXPS (27, 28). We used one of these compounds to co-crystallize with our homologs as a diphosphate-free version of ThDP. The compounds were tested against *paDXPS* and *kpDXPS*. The co-crystallized thiamine analog compound (2-{3-[(4-Amino-2-Methylpyrimidin-5-yl)methyl]phenyl}ethanol) showed a half-maximal inhibitory concentration (IC₅₀) of

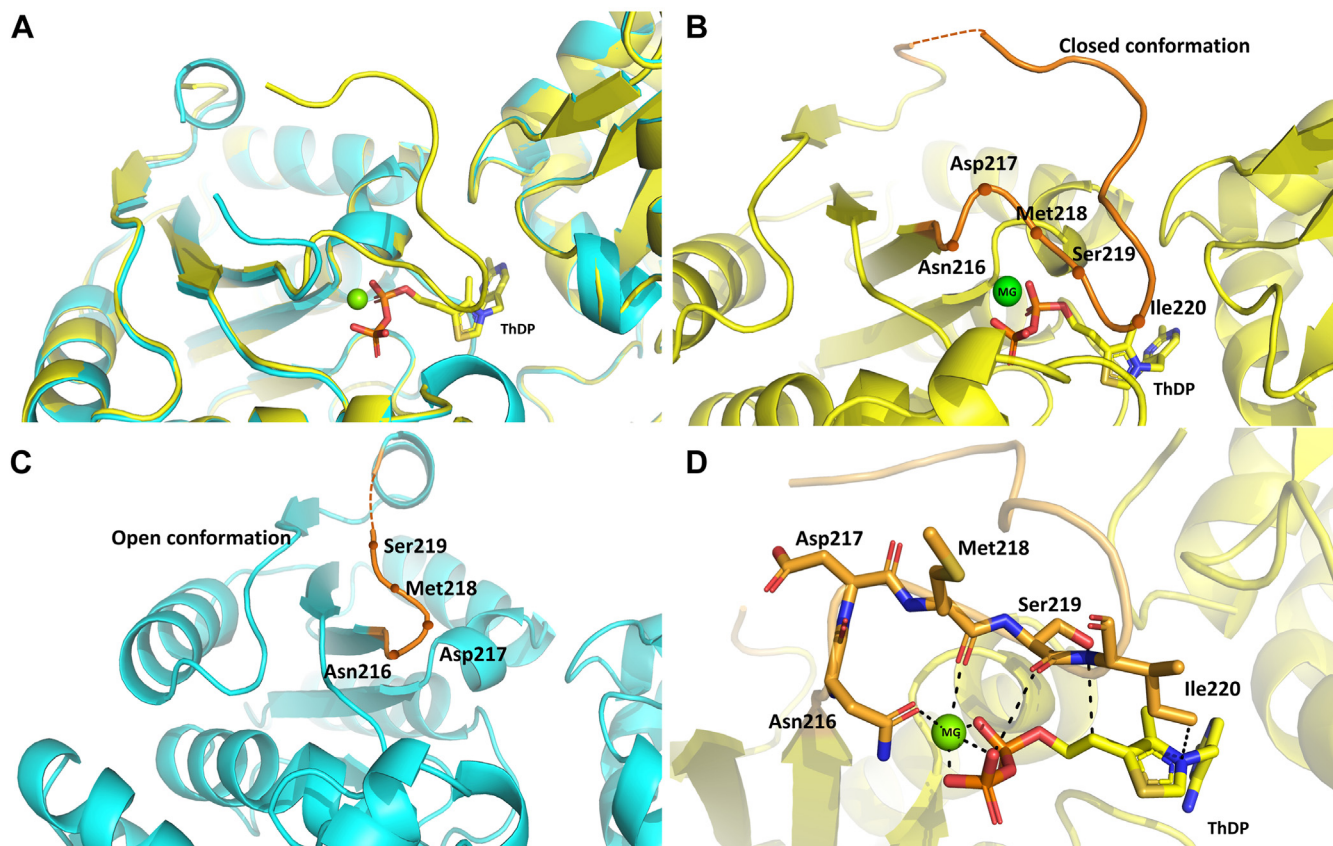


Figure 3. Diphosphate loop movement in the active site of *paDXPS*. A, structure of ThDP-bound *paDXPS* in yellow is superimposed with the structure of *apo paDXPS* in cyan, showing the conformational change of the active site in the bound and unbound state. B, loop in the ThDP bound structure extending from Asn216 to Trp250, in orange exhibiting a closed conformation; C, the open conformation of the same loop. D, highlighting residues, Asn216, Asp217, and Met218, involved in the coordination with the diphosphate. Interactions are shown as black, dashed lines (below 5 Å). All figures of this type were created in PyMol.

DXPS conformational changes upon cofactor binding

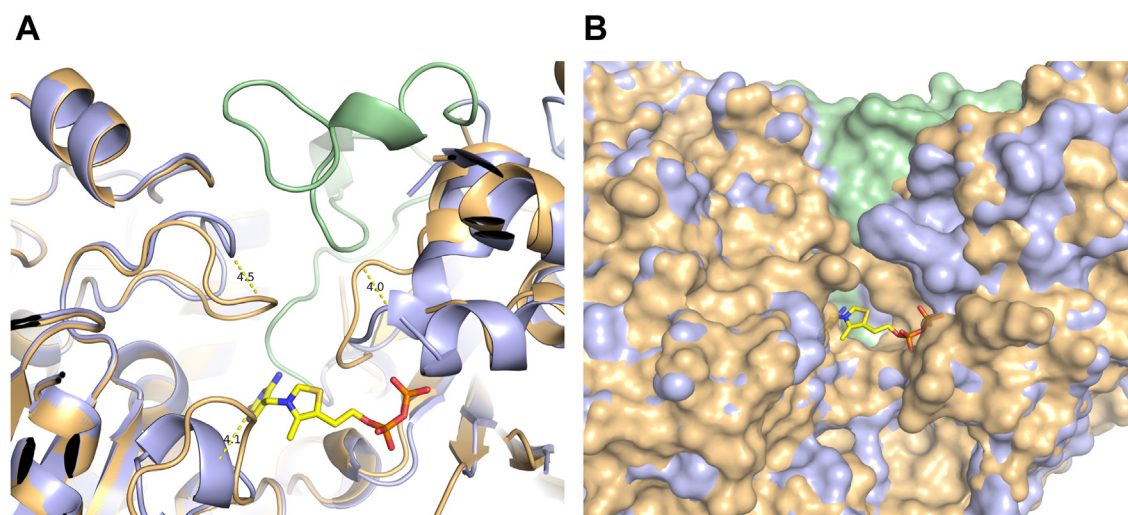


Figure 4. Loops rearrangement close to the active site in *kpDXPS*. *A*, the two versions of *kpDXPS* are visualized as a cartoon representation in *light blue* (ThDP-bound) and *light orange* (*apo*). An N-terminal disordered loop in *apo kpDXPS* (amino acids 91–124) is shown in *pale green*. *B*, comparison of the surface representation of the *kpDXPS* active site, with the same color-coding as in (*A*). Loop movement leads to a much narrower pocket in *apo kpDXPS*, however, the missing N-terminal loop opens up a tunnel on the other side of the cofactor-binding site.

197 μM for *paDXPS* and 94 μM for *kpDXPS*. To further study the conformational changes in the active site loop, we set out to obtain an X-ray structure of either *paDXPS* or *kpDXPS* with the thiamine analog. We were unsuccessful to generate co-crystals for *kpDXPS* with the inhibitor. We were, however, able to obtain a 2.2 Å co-crystal structure of *paDXPS* (PDB ID: 8A4D).

As anticipated, the binding of this analog is similar to the binding of the thiamine moiety of ThDP (pyrimidine and thiazole) with the pyrimidine ring stacking against Phe395. However, the phenyl ring is showing a slight deviation from the plane of thiazole binding seen with ThDP. The hydroxyl group of the ligand is pointing toward the entry site of the substrates while the flexibility of the hydroxyethyl group resulted in poor density for this part of the inhibitor. Nonetheless, it is surrounded by a histidine-rich region and shows the potential of H bonding to these residues (Fig. 5B). Compounds with shorter hydroxyl linkages show less activity (results not shown). Interestingly, the binding of this inhibitor excludes the part of Domain I that is involved in coordinating the diphosphate/Mg²⁺; the diphosphate loop shows the same open conformation as seen in the *apo* enzyme (Fig. 5A). This adds to the notion that this loop is stabilized by the interaction with this specific part of ThDP. This is an especially critical observation for drug discovery, as a different three-dimensional shape of the active site could enable the exploration of different regions of the DXPS in search for inhibitors.

Structure of *paDXPS* in the presence of the substrate analog fluoropyruvate shows the binding of acetylated ThDP and competitive inhibition

Fluoropyruvate is an analog of the natural substrate pyruvate; it was also described as the first inhibitor of DXPS from *P. aeruginosa* (9). DXPS decarboxylates pyruvate in an interaction mediated by ThDP in a stepwise manner: pyruvate first

interacts with ThDP to form the intermediate lactyl ThDP (LThDP), this intermediate was shown to be stabilized by a histidine network near the C2 of ThDP. This reaction promotes a closed conformation of a loop close to the active site and the addition of D-GAP leads to a more open conformation, which eventually encourages decarboxylation (15, 17).

We performed an inhibition assay and a mode-of-inhibition (MOI) study on *paDXPS* using fluoropyruvate. In the inhibition assay, fluoropyruvate inhibited the enzyme with an (IC₅₀) value of around 27 μM . The MOI study indicated competitive inhibition with the natural substrate pyruvate (Table S3). Fluoropyruvate has also been reported to inactivate the E1 subunit of the pyruvate dehydrogenase complex, a mammalian ThDP-dependent enzyme. It was originally proposed that the fluoropyruvate interaction with the ThDP-bound enzyme results in a modification of a sulfhydryl group in the active site (8). Illustrated in Figure 1 is the interaction of fluoropyruvate with ThDP-bound *paDXPS* as compared to pyruvate. To determine whether the inhibition of the enzyme is the result of the acetylation of a reactive residue in the active site after the addition of fluoropyruvate, we performed intact MS measurements to detect the acetylation of the enzyme by mass shift, intact mass results show similar masses for fluoropyruvate treated and untreated proteins (Fig. S5). These results, together with the MOI assay, suggest that no covalent modification is involved in inhibition by fluoropyruvate.

We then opted to get structural information on the fluoropyruvate-treated *paDXPS*. The enzyme was incubated and crystallized with excess fluoropyruvate in the presence of ThDP and MgCl₂. The structure was solved using molecular replacement and refined to 2.0 Å (PDB ID: 8A45). We found no evidence of an acetylated residue within the active site. On the other hand, we observed convincing electron density for the reaction intermediate 2-acetyl-ThDP in the active site (Fig. 5), consistent with the proposed reaction mechanism (Fig. 1). These findings indicate that inhibition with

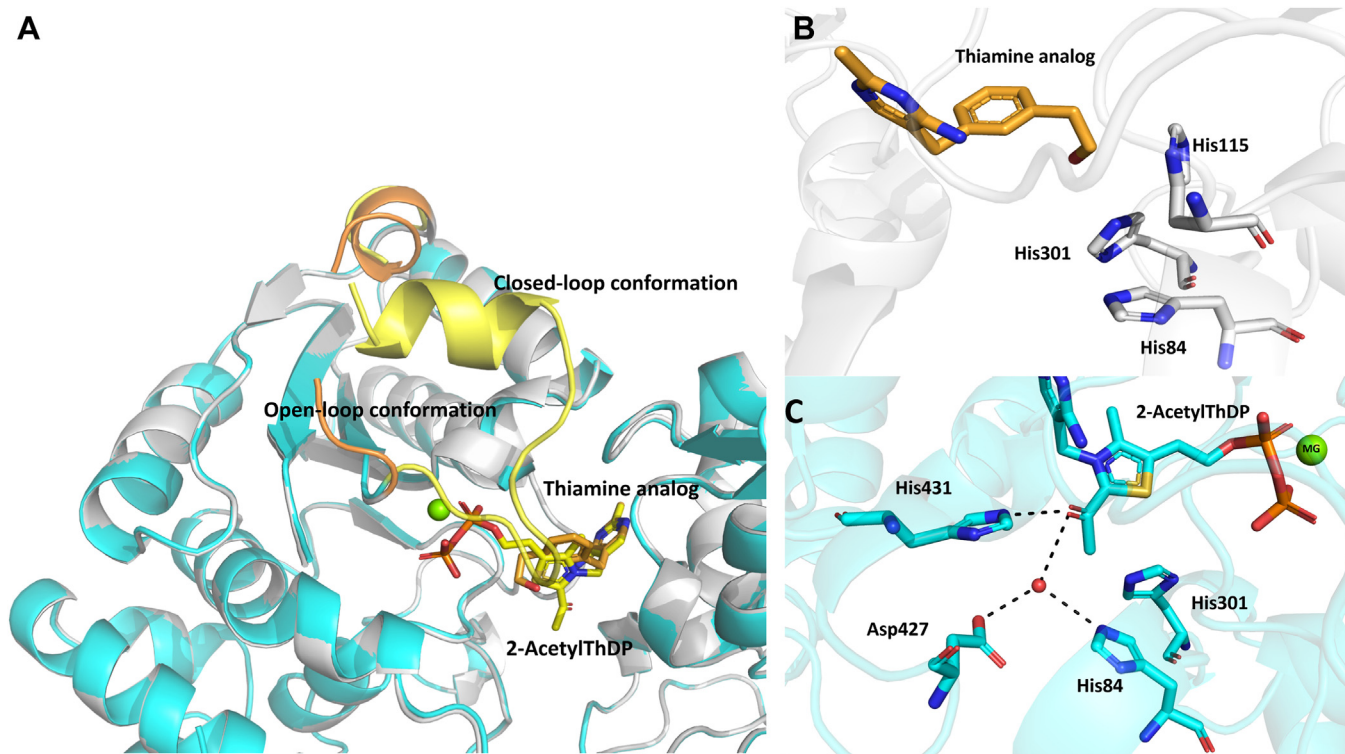


Figure 5. Binding of the thiamine analog and 2-acetyl-ThDP with *paDXPS*. A, thiamine analog bound structure in gray (loop in orange) and 2-acetyl ThDP structure in cyan (loop in yellow). B, zoomed in view of the interactions between thiamine analog (orange) and binding residues of *paDXPS*, the structure shows the hydroxyethyl part of the ligand pointing towards a histidine-rich region in the substrate binding site. The diphosphate loop shows an open conformation in the thiamine-bound structure, in orange. C, interactions between 2-acetyl-ThDP (cyan) and binding residues of *paDXPS*, the structure shows a closed diphosphate loop conformation, in yellow. Interactions are shown as black, dashed lines (below 5 Å).

fluoropyruvate results rather from the acetylated ThDP, which is expected to be less reactive than ThDP and competes with the reactive enamine formed with pyruvate.

The structure of *paDXPS* in the presence of fluoropyruvate provides a high-resolution view of acetylThDP-bound enzyme, a reaction product related to the enamine intermediate. The pyrimidine and diphosphate moiety of 2-acetyl-ThDP is bound in a highly similar way to ThDP (Fig. 5A), with the diphosphate loop also stabilized by coordinating residues Asp187, Asn216, and Met218 to the diphosphate/Mg part. Hydrogen bonds are formed between His-431 and the carbonyl of acetyl-ThDP. This histidine residue is conservatively preserved in other ThDP-dependent enzymes and has been shown to contribute to a relaxed/packed state, that ultimately drives decarboxylation (29). The conformation of this region is different in the structure of *mtDXPS*. *M. tuberculosis* replaces this interaction with a network of water-coordinated histidine, serine, and tyrosine (23). Interestingly, the carbonyl is also involved in coordinating water molecules with Asp427 and His84, while preserving the his431 conformation (Fig. 5C). The presence of these interactions results in a tetrahedral geometry around C2 of 2-acetylThDP, causing it to deviate from the plane of the thiazole, enabled by the single bond in the acetyl-ThDP compared to the double bond in the enamine. Our structure also shows the presence of His301, adding a positive charge to the pocket, which is missing in the structure with enamine in *drDXPS*. This histidine is replaced by tyrosine in *ecDXPS* (14). His301 is also part of the spoon-fork motif, the flexibility of

this site facilitates binding to the substrates and stabilizes intermediates (17). This region is significantly different from the DXPS structures. Figure 6 shows only some differences within the aligned fork motifs of the 2-acetyl ThDP *paDXPS* structure and *drDXPS* structure. In *kpDXPS* residues 246 to 284 are disordered and are also part of the spoon-fork motif. These differences might result in even more different conformational changes in this region during catalysis.

Conclusion

This study introduces the first *apo* structures of DXPS from *P. aeruginosa* and *K. pneumoniae*, along with their cofactor ThDP-bound structures. These findings contribute to expanding our understanding of this important enzyme class and provide crucial structural insights for the advancement of innovative drug development targeting ESKAPE pathogens.

Upon ThDP binding, DXPS undergoes conformational changes that lead to an open and closed loop conformation near the active site. In the *paDXPS* structure, the closed conformation is stabilized by interactions of the ThDP diphosphate/Mg²⁺ moiety with an asparagine-rich region of near the active site. In contrast, an open conformation is seen when this interaction is lost in the *apo* structure. This hypothesis was also confirmed by analyzing the crystal structure of *paDXPS* with a thiamine analog lacking the diphosphate moiety, in which the same loop adopts the open conformation seen with the *apo* enzyme. The DXPS structure from

DXPS conformational changes upon cofactor binding

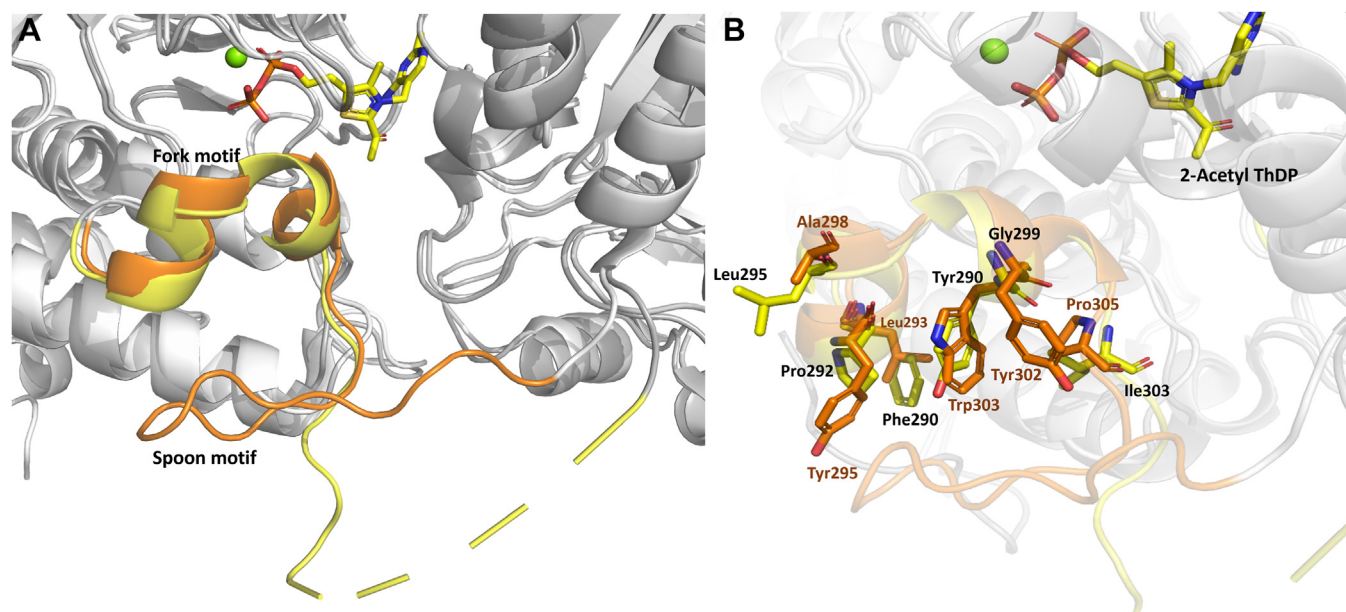


Figure 6. Spoon-fork motif of *paDXPS*. A, spoon-fork motif in *paDXPS* in the presence of 2-acetyl ThDP colored yellow, superimposed with *drDXPS* in the presence of PLThDP, colored in orange showing the open conformation of the spoon motif in both structures. B, zoomed-in view of the fork motif highlighting changes in residues involved in stabilizing conformational changes in this region. Residues from *drDXPS* are labeled in orange, from *paDXPS* in black.

K. pneumoniae also shows loop movement deep in the active site, which leads to a much narrower pocket in the absence of ThDP.

From a drug-design perspective, these results open the door to different modes of DXPS inhibition. Our structures reveal the important role of the diphosphate moiety of ThDP in stabilizing part of the active site. Targeting specific residues in this loop may prevent the enzyme from binding to ThDP by not providing enough space for ThDP to bind, especially with the rearrangement of the loops in *kpDXPS*. In *paDXPS*, one strategy could be to prevent the diphosphate-binding loop from closing, thereby preventing the enzyme from entering its active conformation. This could perhaps be achieved by using the pyrimidyl ring present in ThDP as a warhead and optimizing the rest of the compound to bind to the loop in its “open” conformation. Additionally, the binding mode of the thiamine analog indicates that it is also possible to explore compounds that target the substrate-binding channel in addition to the ThDP binding pocket. Following this approach, a three-headed inhibitor could perhaps be developed, which would likely show increased selectivity for DXPS over other ThDP-dependent enzymes due to the exploitation of DXPS’s unique binding-site geometry. Furthermore, we demonstrated that a competitive inhibition could be achieved by blocking further interactions of ThDP-bound enzyme with molecules that can form direct, less reactive intermediates with ThDP, therefore trapping the enzyme in this rate-determining kinetic step and preventing subsequent interactions with the natural substrates. These additional DXPS X-ray crystal structures from important ESKAPE pathogens lay the foundation for future structure-based drug design of potent and selective small-molecule inhibitors with a novel mode of action to treat bacterial infections.

Experimental procedures

Protein expression and purification

paDXPS

The construct genes were synthesized commercially, cloned into pET28a, and transformed into BL21(DE3) for over-expression. The transformed cells were grown in a selective LB medium supplemented with Kanamycin (50 mg/ml) up to an OD₆₀₀ of 0.6, then induced with 1 mM isopropyl-L-D-thiogalactoside (IPTG) and were subsequently grown at 18 °C for 16 h. The harvested cells were then disrupted in a microfluidizer after resuspension in a washing buffer consisting of 50 mM HEPES, 100 mM NaCl, 20 mM imidazole, and 2 mM β-mercaptoethanol (100 ml buffer per 25 g of wet cell pellet). After centrifugation, the supernatant was loaded onto a 5 ml HisTrap HP column equilibrated in washing buffer. After an extensive washing step with 30 column volumes of washing buffer, the bound fraction was then eluted with 50 mM HEPES, 100 mM NaCl, 300 mM Imidazole, and 2 mM β-mercaptoethanol. His-tag purification was followed by size-exclusion chromatography in 20 mM HEPES, 200 mM NaCl, and 2 mM β-mercaptoethanol. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and protein mass spectrometry were used to confirm the purity of protein purifications.

kpDXPS

The construct gene was synthesized commercially, cloned into pET28b+, and transformed into Lemo21(DE3) for over-expression. The transformed cells were grown in a selective LB medium at 37 °C supplemented with kanamycin (50 mg/ml) and chloramphenicol (34 mg/ml) up to an OD₆₀₀ of 0.9, then induced with 0.1 mM isopropyl-L-D-thiogalactoside (IPTG), and were subsequently grown at 18 °C for 16 h. The harvested cells were

then disrupted in a microfluidizer after resuspension in a washing buffer consisting of 20 mM Tris pH 8.0, 200 mM NaCl, 20 mM Imidazole pH 8.0, and 5% glycerol (100 ml buffer per 25 g of wet cell pellet). After centrifugation, the supernatant was loaded onto a 5 ml HisTrap HP column equilibrated in washing buffer. After an extensive washing step with 30 column volumes of washing buffer, the bound fraction was then eluted with an elution buffer consisting of 20 mM Tris pH 8.0, 200 mM NaCl, 250 mM imidazole pH 8.0, and 5% glycerol. Eluted protein was passed over a HiPrep 26/10 desalting column equilibrated in washing buffer to remove excess imidazole, after which TEV digestion was carried out by adding TEV protease in a 1:10 protein: TEV ratio and incubating the solution at 4 °C for 16 h. The N-terminal 6x His-Tag was removed by passing the protein over a 5 ml His-Trap HP column and collection of the flow-through. Purification was finalized by size-exclusion chromatography in 20 mM Tris, 200 mM NaCl, final pH 8.0 for *apo kpDXPS* and 10 mM HEPES, 200 mM NaCl, 5% glycerol, final pH 7.4 for ThDP-bound *kpDXPS*. SDS-PAGE and protein mass spectrometry were used to confirm the purity of protein purifications.

Crystallization

paDXPS

Crystallization trials were carried out with a protein solution containing 15 mg/ml *paDXPS* in a storage buffer using commercially available screens from NeXtalbiotech. Ligand-bound structures were obtained after incubating *paDXPS* with an excess amount of the ligand for 2 h at RT. In every case, crystals were grown in hanging drops with a reservoir solution consisting of 100 mM HEPES, 12% PEG-8000, and 200 mM calcium acetate. Plates were observed every 3 days and optimization of the identified conditions was carried out. 32% Glycerol was added as a cryo-protectant before the crystals were mounted in a cryo-loop and stored in liquid nitrogen until data collection.

kpDXPS

Crystallization trials were carried out with a protein solution containing 12.5 mg/ml *kpDXPS* (ligand-bound) and 47.5 mg/ml *kpDXPS* (*apo*) in the respective storage buffer using commercially available screens from NeXtalbiotech. Ligand-bound structures of *kpDXPS* were obtained by the addition of 1 mM ThDP/MgCl₂ and incubating the solution for 4 h on ice before setting up the crystallization plates. For the protein complex, crystals were grown in sitting-drop optimization plates at 18 °C with a reservoir solution consisting of 100 mM HEPES pH 7.0 to 7.5 and 18 to 28% PEG-3000. For the *apo* version, crystals were grown in hanging-drop optimization plates at 18 °C with a reservoir solution consisting of 200 to 400 mM MgCl₂, 100 mM HEPES pH 7.5, and 14 to 18% PEG-8000. In both cases, 15% 2R,3R-butanediol was added as a cryoprotectant before the crystals were mounted in a cryo-loop and stored in liquid nitrogen until data collection.

Data collection and refinement

X-ray diffraction data for DXPS structures were collected at the Swiss Light Source located at the Paul Scherrer Institute in

Switzerland. Specific information about each beamline can be found in the specific sections, all data collection and refinement statistics can be found in Table S2. Data reduction and scaling were done using AIMLESS (30) in CCP4i (31). Structures were solved using molecular replacement using the program Phaser.MR (32) in Phenix software (33). Repetitive cycles of refinement were done in COOT (34) and Phenix refine (35) to obtain final structures (Tables S1 and S2). Structures were validated using the Molprobtity server and all figures were rendered using PyMoL.

Kinetics and inhibition assays

To measure DXPS activity and kinetic parameters, a coupled spectrophotometric enzyme assay was adapted from the assay protocol described in the literature (36). A continuous kinetic photometric assay was used to measure DXPS activity. NADPH consumption by the auxiliary enzyme 1-deoxy-D-xylulose 5-phosphate reductoisomerase (IspC) was measured in a microplate reader (PHERAstar), monitoring the absorbance at 340 nm. In all cases (IspC) concentration was kept in excess (1 μM) to ensure DXPS is the only rate-limiting enzyme in the reaction mix. In the inhibition assays, in the case of the thiamine analog inhibitor, ThDP concentration was kept around 4X K_m (400 nM) to ensure the enzyme reaches maximum velocity while allowing competitive inhibition. In the case of fluoropyruvate, pyruvate concentration was also kept at 4X K_m (500 μM) for the same reason. For inhibition studies, DXPS activity was analyzed at RT for 30 min as follows; reaction mixture A (100 mM HEPES pH 7.5, 100 mM NaCl, 1 mM MgCl₂, 0.5 μM NAPDH, 1 μM IspC, 400 nM ThDP, and 0.15 μM and 0.2 μM *paDXPS* and *kpDXPS* respectively) was pre-incubated with different concentrations of inhibitor for 5 min at 37 °C in 10% DMSO. The reaction was then started by adding mixture B (100 mM HEPES pH 7.5, 100 mM NaCl, 2 mM DL-glyceraldehyde-3-phosphate and 2 mM pyruvate (thiamine analog inhibitor) and 500 μM pyruvate (fluoropyruvate)). For kinetic assays, the cofactor ThDP and substrates pyruvate and D-GAP were used in varying concentrations. When a substrate or cofactor was kept constant, a concentration of 400 nM was used for ThDP and 2 mM for pyruvate and D-GAP. Blank correction and linear fitting of the absorption data were performed using the program Origin 2019 (Origin-Lab). The initial velocities obtained were plotted against the substrate concentrations, and the K_m values were determined by nonlinear curve fitting using the Michaelis–Menten model of the enzyme kinetics, Figs. S6 and S7.

Data availability

Data supporting this study are included within the article and/or supporting materials.

Supporting information—This article contains supporting information.

Acknowledgments—The authors would like to thank Dr Nicolas Frank and Dr Timo Risch for native MS and intact MS

DXPS conformational changes upon cofactor binding

measurements. We acknowledge the Paul Scherrer Institute, Villigen, Switzerland for the provision of synchrotron radiation beamtime at beamline X06DA-PXIII and X06SA-PXI of the SLS and the European Synchrotron Radiation Facility (ESRF), Grenoble, France for provision of beamline ID23 to 1.

Author contributions—A. K. H. H., R. H., and S. A. conceptualization; R. H. and S. A. writing—original draft; A. L. and L. M. synthesis. A. K. H. H., R. H., S. A., A. L., L. M., and J. K., writing—review and editing; A. K. H. H. supervision.

Funding and additional information—The Helmholtz Association's Initiative and Networking Fund (to A. K. H. H.) and the Schlumberger Foundation faculty for the future (FFTF) (to R. H.) funded this work.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: D-GAP, glyceraldehyde 3-phosphate; DXP, D-xylulose-5-phosphate; DXPS, D-xylulose-5-phosphate synthase; LthDP, lactyl thiamine diphosphate; MEP, methylerythritol-phosphate; *paDXPS*, DXPS from *P. aeruginosa*; *kpDXPS*, *K. pneumoniae* DXPS; ThDP, thiamine diphosphate.

References

- Lin, J., Nishino, K., Roberts, M. C., Tolmasky, M., Aminov, R. I., and Zhang, L. (2015) Mechanisms of antibiotic resistance. *Front. Microbiol.* **6**, 34
- Miethke, M., Pieroni, M., Weber, T., Brönstrup, M., Hammann, P., Halby, L., *et al.* (2021) Towards the sustainable discovery and development of new antibiotics. *Nat. Rev. Chem.* **5**, 726–749
- Nolan, L. M., Turnbull, L., Katrib, M., Osvath, S. R., Losa, D., Lazenby, J. J., *et al.* (2020) *Pseudomonas aeruginosa* is capable of natural transformation in biofilms. *Microbiology (Reading)* **166**, 995–1003
- Camus, L., Briaud, P., Vandenesch, F., and Moreau, K. (2021) How bacterial adaptation to cystic fibrosis environment shapes interactions between *Pseudomonas aeruginosa* and *Staphylococcus aureus*. *Front. Microbiol.* **12**, 617784
- Riquelme, S. A., Ahn, D., and Prince, A. (2018) *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* adaptation to innate immune clearance mechanisms in the lung. *J. Innate Immun.* **10**, 442–454
- Magill, S. S., Edwards, J. R., Bamberg, W., Beldavs, Z. G., Dumyati, G., Kainer, M. A., *et al.* (2014) Multistate point-prevalence survey of health care-associated infections. *N. Engl. J. Med.* **370**, 1198–1208
- Rodriguez-Concepcion, M. (2004) The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *CPD* **10**, 2391–2400
- Flournoy, D. S., and Frey, P. A. (1989) Inactivation of the pyruvate dehydrogenase complex of *Escherichia coli* by fluoropyruvate. *Biochemistry* **28**, 9594–9602
- Altincicek, B., Hintz, M., Sanderbrand, S., Wiesner, J., Beck, E., and Jomaa, H. (2000) Tools for discovery of inhibitors of the 1-deoxy-D-xylulose 5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with enzymes from the pathogenic bacterium *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* **190**, 329–333
- Gräwert, T., Groll, M., Rohdich, F., Bacher, A., and Eisenreich, W. (2011) Biochemistry of the non-mevalonate isoprenoid pathway. *Cell. Mol. Life Sci.* **68**, 3797–3814
- Wang, X., and Dowd, C. S. (2018) The methylerythritol phosphate pathway: promising drug targets in the Fight against tuberculosis. *ACS Infect. Dis.* **4**, 278–290
- Smith, J. M., Vierling, R. J., and Meyers, C. F. (2012) Selective inhibition of *E. coli* 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *Med. Chem. Commun.* **3**, 65–67
- Smith, J. M., Warrington, N. V., Vierling, R. J., Kuhn, M. L., Anderson, W. F., Koppisch, A. T., *et al.* (2014) Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *J. Antibiot. (Tokyo)* **67**, 77–83
- Xiang, S., Usunow, G., Lange, G., Busch, M., and Tong, L. (2007) Crystal structure of 1-Deoxy-d-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J. Biol. Chem.* **282**, 2676–2682
- Patel, H., Nemeria, N. S., Brammer, L. A., Freil Meyers, C. L., and Jordan, F. (2012) Observation of thiamin-bound intermediates and microscopic rate constants for their Interconversion on 1-deoxy- d -xylulose 5-phosphate synthase: 600-fold rate Acceleration of pyruvate decarboxylation by d -Glyceraldehyde-3-phosphate. *J. Am. Chem. Soc.* **134**, 18374–18379
- Bartee, D., and Freil Meyers, C. L. (2018) Toward understanding the chemistry and biology of 1-deoxy- d -xylulose 5-phosphate (DXP) synthase: a unique antimicrobial target at the Heart of bacterial metabolism. *Acc. Chem. Res.* **51**, 2546–2555
- Chen, P. Y.-T., DeColli, A. A., Freil Meyers, C. L., and Drennan, C. L. (2019) X-ray crystallography-based structural elucidation of enzyme-bound intermediates along the 1-deoxy-d-xylulose 5-phosphate synthase reaction coordinate. *J. Biol. Chem.* **294**, 12405–12414
- Brammer Basta, L. A., Patel, H., Kakalis, L., Jordan, F., and Freil Meyers, C. L. (2014) Defining critical residues for substrate binding to 1-deoxy-d-xylulose 5-phosphate synthase – active site substitutions stabilize the predecarboxylation intermediate C2 α -lactylthiamin diphosphate. *FEBS J.* **281**, 2820–2837
- Zhou, J., Yang, L., DeColli, A., Freil Meyers, C., Nemeria, N. S., and Jordan, F. (2017) Conformational dynamics of 1-deoxy-D-xylulose 5-phosphate synthase on ligand binding revealed by H/D exchange MS. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9355–9360
- Bartee, D., and Freil Meyers, C. L. (2018) Targeting the unique mechanism of bacterial 1-deoxy-D-xylulose 5-phosphate (DXP) synthase. *Biochemistry* **57**, 4349–4356
- Gierse, R. M., Redeem, E., Diamanti, E., Wrenger, C., Groves, M. R., and Hirsch, A. K. (2017) DXS as a target for structure-based drug design. *Future Med. Chem.* **9**, 1277–1294
- Masini, T., Kroezen, B. S., and Hirsch, A. K. H. (2013) Druggability of the enzymes of the non-mevalonate-pathway. *Drug Discov. Today* **18**, 1256–1262
- Gierse, R. M., Oerlemans, R., Reddem, E. R., Gawriljuk, V. O., Alhayek, A., Baitinger, D., *et al.* (2022) First crystal structures of 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) from *Mycobacterium tuberculosis* indicate a distinct mechanism of intermediate stabilization. *Sci. Rep.* **12**, 7221
- Gierse, R. M., Reddem, E. R., Alhayek, A., Baitinger, D., Hamid, Z., Jakobi, H., *et al.* (2021) Identification of a 1-deoxy-D-xylulose-5-phosphate synthase (DXS) mutant with improved crystallographic properties. *Biochem. Biophys. Res. Commun.* **539**, 42–47
- Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
- Versées, W., Spaepen, S., Wood, M. D. H., Leeper, F. J., Vanderleyden, J., and Steyaert, J. (2007) Molecular mechanism of allosteric substrate activation in a thiamine diphosphate-dependent Decarboxylase. *J. Biol. Chem.* **282**, 35269–35278
- Mann, S., Perez Melero, C., Hawksley, D., and Leeper, F. J. (2004) Inhibition of thiamin diphosphate dependent enzymes by 3-deazathiamin diphosphate. *Org. Biomol. Chem.* **2**, 1732
- Masini, T., Lacy, B., Monjas, L., Hawksley, D., de Voogd, A. R., Illarionov, B., *et al.* (2015) Validation of a homology model of *Mycobacterium tuberculosis* DXS: rationalization of observed activities of thiamine derivatives as potent inhibitors of two orthologues of DXS. *Org. Biomol. Chem.* **13**, 11263–11277
- Arjunan, P., Sax, M., Brunskill, A., Chandrasekhar, K., Nemeria, N., Zhang, S., *et al.* (2006) A thiamin-bound, pre-decarboxylation reaction intermediate analogue in the pyruvate dehydrogenase E1 subunit induces large scale disorder-to-order transformations in the enzyme and reveals novel structural features in the covalently bound adduct. *J. Biol. Chem.* **281**, 15296–15303

30. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214
31. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
32. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674
33. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
34. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
35. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
36. Takahashi, S., Kuzuyama, T., Watanabe, H., and Seto, H. (1998) A 1-deoxy-d-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-d-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9879–9884

Supplementary:**1-deoxy-D-xylulose 5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding****Rawia Hamid^{1, 2}, Sebastian Adam¹, Antoine Lacour¹, Leticia Monjas³, Jesko Köhnke⁴, Anna K. H. Hirsch^{1, 2*}.**

1. Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123 Saarbrücken, Germany.

2. Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany.

3- Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, NL-9747 AG Groningen, the Netherlands.

4- Institute of Food Chemistry, Leibniz University Hannover, Callinstr. 5 30167 Hannover, Germany. School of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK.

*Correspondence e-mail: anna.hirsch@helmholtz-hips.de

Table S1: Primary amino acid sequences of proteins used in this study.	32
Table S2: Data collection and refinement statistics	33
Figure S1: Sequence alignment of DXPS homologs	34
Figure S2: Native Ms analysis of <i>paDXPS</i>	35
Figure S3: Calibration curve used to estimate molecular weights for DXPS	35
Figure S4: DXPS homologs active site comparison	36
Figure S5: Mass spectra of <i>paDXPS</i> before and after incubation with fluoropyruvate	37
Figure S6: Michaelis-Menten kinetics analysis of <i>paDXPS</i> and <i>kpDXPS</i> enzymes	38
Figure S7: Inhibitory dose-response curves IC ₅₀ of inhibitor on <i>paDXPS</i> and <i>kpDXPS</i>	38
Table S3. Determination of MOI of fluoropyruvate against <i>paDXPS</i>	39

Table S1: Primary amino acid sequences of proteins used in this study.

Construct	Primary amino acid sequence
Native paDXPS	MGSSHHHHHSSGLVPRGSMENLYFQSHMPKTLHEIPRERPATPLLDRASSPAELRRLGEADLE TLADELRQYLLYTVGQTGGHFGAGLGVVELTIALHYVFDTPDDRLLVWDVGHQAYPHKILTERREL MGTLRQKNGLAAFPRAESEYDTFGVGHSSSTSISAALGMAIAARLQGKERKSVAVIGDGALTAG MAFEALNHASEVDADMLVILNDNDMSISHNVGGLSNYLAKILSSRTYSSMREGSKKVLRLPGA WEIARRTEEYAKGMLVPGTLFEELGWNYIGPIDGHDLPTLVATLRNMRDMKGPQFLHVVTKKGK GFAPAELDPIGYHAITKLEAPGSAPKKTGGPKYSSVFGQWLCDMAAQDARLLGITPAMKEGSDL VAFSERYPERYFDVAIAEQHAVTLAAGMACEGMKPVVAIYSTFLQRAYDQLIHDVAVQHLDVLFAL DRAGLVGEDGPTHAGSFDISYLRCPGMLVMTSPDEDELKLLTTGYLFDGPAAVRYPRGSGPN HPIDPDLQPVEIGKGVVRRRGRRVALLVFGVQLAEAMKVAESLDATVVDMRFVKPLDEALVRELA GSHELLVTIEENAVMGGAGSAVGEFLASEGLEVPDLLQLGLPDYVEHAKPSEMLAECGLDAAGIE KAVRQLDRQ
Mutated paDXPS	MGSSHHHHHSSGLVPRGSMENLYFQSHMPKTLHEIPRERPATPLLDRASSPAELRRLGEADLE TLADELRQYLLYTVGQTGGHFGAGLGVVELTIALHYVFDTPDDRLLVWDVGHQAYPHKILTERREL MGTLRQKNGLAAFPRAESEYDTFGVGHSSSTSISAALGMAIAARLQGKERKSVAVIGDGALTAG MAFEALNHASEVDADMLVILNDNDMSISHNVGGLSNYLAKIGGGGGGPGTLFEELGWNYIGPIDG HDLPTLVATLRNMRDMKGPQFLHVVTKKGKGFAPAELDPIGYHAITKLEAPGSAPKKTGGPKYSS VFGQWLCDMAAQDARLLGITPAMKEGSDLVAFSERYPERYFDVAIAEQHAVTLAAGMACEGMK PVVAIYSTFLQRAYDQLIHDVAVQHLDVLFALDRAGLVGEDGPTHAGSFDISYLRCPGMLVMTSP DEDELKLLTTGYLFDGPAAVRYPRGSGPNHPIDPDLQPVEIGKGVVRRRGRRVALLVFGVQLA EAMKVAESLDATVVDMRFVKPLDEALVRELAGSHELLVTIEENAVMGGAGSAVGEFLASEGLEVP DLLQLGLPDYVEHAKPSEMLAECGLDAAGIEKAVRQLDRQ
Mutated kpDXPS	MSFDIAKYPTLALVDSTQELRLLPKESLPKLCDELRRYLLDSVSRSSGHFASGLGTVELTVALHYV YNTPFDRLIWDVGHQAYPHKILTGRRDKIGTIRQKGGGLHPPFWRGESEYDVLVSVGHSSSTSISAGI GVAIAAAKEDKQRRAVCVIGDGAITAGMAFEAMNHAGDIKPDLLVVLNDNEMSISENVGALNNHL AGGGGGGPGTLFEELGFNYIGPVDGHDVGLVSTLKNMRDLKGPQFLHIMTKKGRGYEPAEK DPITFHAVPKFDHTSGVLPKSSGGLPSYSKIFGDWLCETAADKNKLMAITPAMREGSGMVEFSKK FPDYFDVAIAEQHAVTFAAGLAIGDYKPVVAIYSTFLQRAYDQVIHDVAIQKLPVLFIDRAGIVGA DGQTHQGAFDLSFLRCIPDMVVMTPSDENECRQMLYTGYYHSDGPCAVRYPRGSGTGATLEPL ASLPIGKGVVQRQGEKIAILNFGTLLPEAAAVADKLNATLVDMRFVKPLDTALILQLAGEHDALVTL EENAIMGGAGSGVNEVLMARRAVPVLNIGLDPDYFIPQGTQEEIRADLGLDAAGIEAKIRDWLA

Table S2: Data collection and refinement statistics of all X-ray structures shown in the manuscript. Statistics for the highest-resolution shell are shown in parentheses.

	Apo paDXPS	ThDP-bound paDXPS	paDXPS – 2-acetyl-ThDP	paDXPS – Thiamine analog	ThDP-bound kpDXPS	Apo kpDXPS
PDB code	8A29	8A5K	8A45	8A4D	8A9C	8A8Y
Resolution range	48.59–2.2	46.29–2.37	48.59–2.0	48.54–2.2	43.37 – 1.8	91.87 – 2.1
Space group	P21 21 21	P21 21 21	P21 21 21	P21 21 21	P21	C2 2 21
Unit cell a b c [Å] $\alpha \beta \gamma$ [°]	116.44 137.63 232.08 90.00 90.00 90.00	116.82Å 137.06Å 231.44 90 90 90	117.12 138.01 231.63 90.00 90.00 90.00	116.618 137.618 231.681 90.00 90.00 90.00	89.67 72.56 91.44 90.00 108.46 90.00	119.16 144.25 142.91 90.0 90.0 90.0
Total reflections	376678 (37313)	325476 (32403)	501348 (49688)	371083 (36366)	557529 (75397)	323919 (45609)
Unique reflections	188736 (18663)	164247 (16316)	187627 (18663)	188257 (18546)	100675 (14606)	71565 (10344)
Multiplicity	7.1 (7.3)	2.0 (2.0)	12.2 (12.3)	4.5 (4.5)	5.5 (5.2)	4.5 (4.4)
Wavelength	0.97	0.97	0.97	0.97	0.97	0.97
Completeness (%)	99.9 (100)	99.46 (99.83)	99.34 (99.50)	99.63	97.9 (97.6)	99.6 (99.6)
Mean I/ sigma (I)	11.09 (2.65)	6.52 (2.48)	14.56 (3.47)	6.86 (1.65)	13.4 (2.5)	9.9 (2.2)
Wilson B-factor	19.66	25.62	19.67	25.24	25.41	35.38
R-merge	0.140 (0.953)	0.07279 (0.3303)	0.150 (1.187)	0.1092 (0.5661)	0.066 (0.561)	0.089 (0.720)
R-work	0.1896	0.1825	0.1948	0.1996	0.1593	0.1749
R-free	0.2254	0. 238	0.2261	0.2323	0.1902	0.1987
Number of non-hydrogen atoms	27762	27348	28232	28005	8837	7781
Macromolecules	25575	25645	25744	25672	7984	7384
Ligands	70	280	296	213	26	
Solvent	2157	1527	2300	2222	827	397
Protein residues	3357	3370	3384	3366	1043	970
RMS (bonds)	0.003	0.019	0.005	0.002	0.006	0.003
RMS (angles)	0.63	0.93	0.70	0.55	0.79	0.56
Average B-factor	25.78	30.88	29.70	31.57	32.6	48.29
Macromolecules	26.85	30.86	29.30	29.40	31.84	48.42

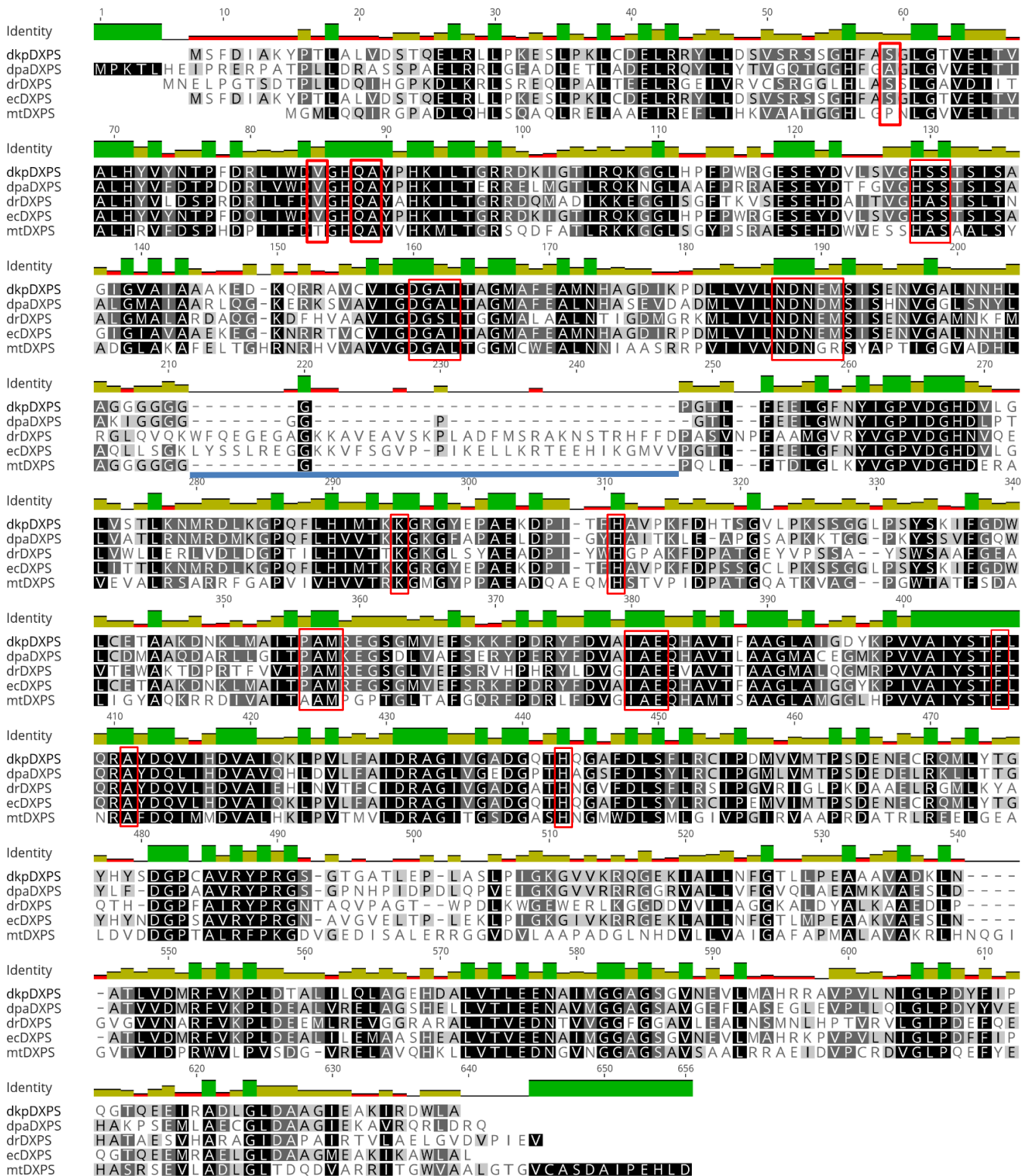


Figure S1: Sequence alignment of DXPS homologs: from *p. aeruginosa*, *K. pneumoniae*, *d. radiodurans* and *E. coli*. The identity, shown as a bar graph above the sequences, was calculated using the four aligned sequences. Amino acids in the active site are indicated by red boxes. The truncated loop is underlined with bold blue line.

Native Mass Spectrometry

Purified *paDXPS* was buffer-exchanged into 10 mM ammonium acetate solution (pH 7.0) by repeated ultracentrifugation using 10K centrifugal filters. The protein was diluted to 5 μ M to obtain the highest sensitivity in the mass spectrometer. Native MS analysis was conducted on a Bruker Solarix XR 7T Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer equipped with a high-resolution electrospray ionization (HRESI) source (Bruker Daltonics, Billerica, MA, USA) by direct infusion of the mixture at a flow rate of 2 μ L/min. Source parameters were set to 500V end-plate offset, 4000 V capillary voltage, 3 bar nebulizer gas pressure, 5 L/min dry gas, and 200 °C dry gas temperature. The instrument was calibrated using an Agilent ESI-L low-concentration tuning mix (G1969-85000). Profile spectra were recorded in positive ion mode with a mass range from 150 to 5000 m/z. Each spectrum was a sum of 128 transients composed of 512k data points. The pulse sequence control and data acquisition was controlled by ftms control software in a Windows operating system.

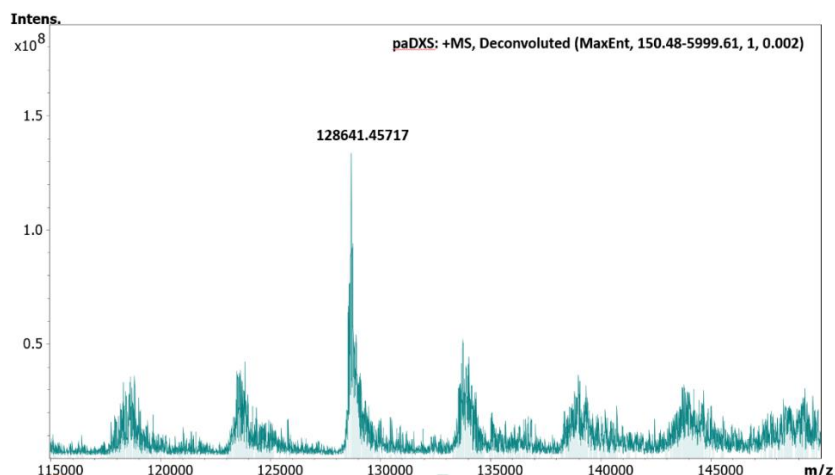


Figure S2: Native Ms analysis of *paDXPS*. The full mass of the *paDXPS* protein can be observed with a m/z of 128641 showing the presence of *paDXPS* as a dimer in solution, the mass corresponds to 2 *paDXPS* monomers (~ 63965 Da), 2 ThDP units. The MS parameters on the Orbitrap EMR were optimized specifically for the m/z window of *paDXPS* protein.

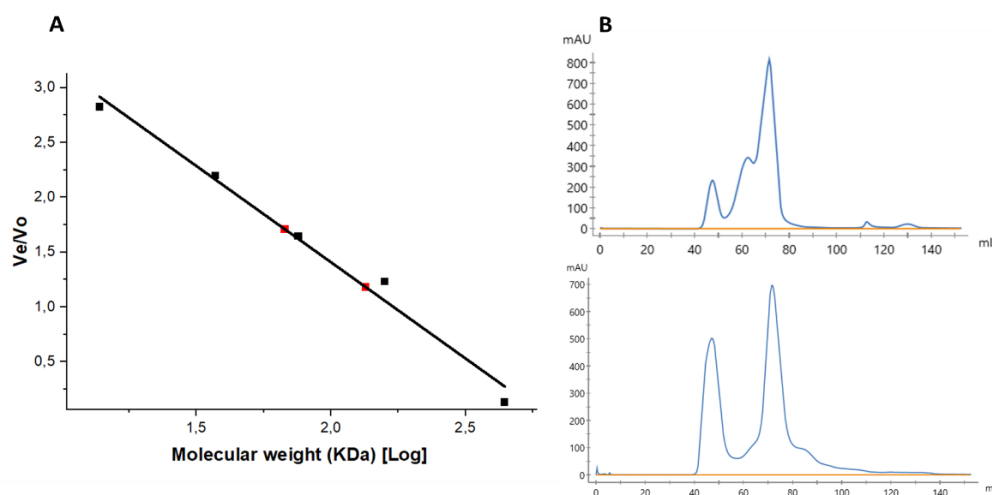


Figure S3: A: Calibration curve used to estimate molecular weights for DXPS. Size exclusion chromatography column S200 16/600 was calibrated using Thyroglobulin (670 KDa), γ -Globulin (158 KDa), ovalbumin (44 kDa), Myoglobin (17 KDa), and Vitamin B12 (1,350 KDa). The calibration curve was plotted using the elution volume/ column void volume (V_e/V_0) versus logarithm of the molecular weight. Straight line is the calibration curve calculated from the data for molecular weight standards. Red dots correspond to the positions of V_e/V_0 values for DXPS. Linear equation, from the calibration curve was used to calculate the experimental molecular weights reported in this study. **B:** elution profile of *kpaDXPS* showing elution in two peaks (UV spectra) corresponding to the monomer and the dimer in SEC.

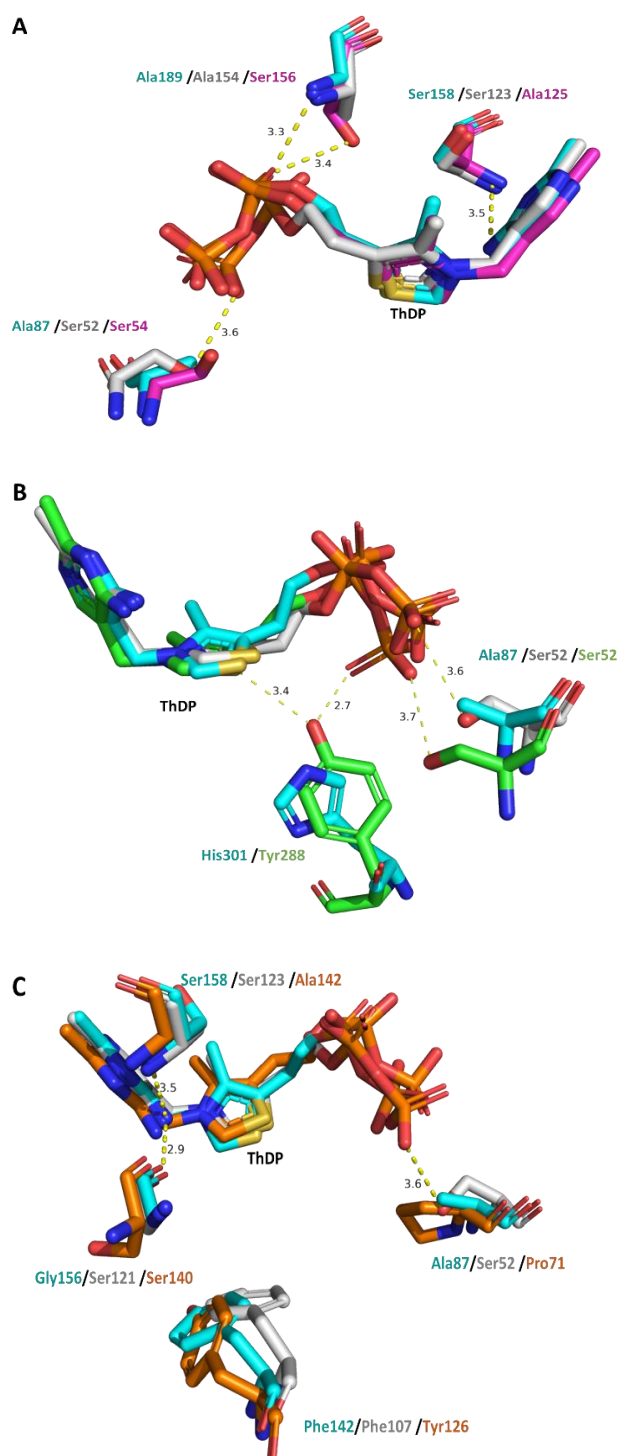


Figure S 4: DXPS homologs active site comparison: showing residues involved in binding to ThDP, highlighting differences in residues in/near the active site. **A** *pa*DXPS in cyan overlaid with *kp*DXPS in grey and the model homolog from *dr*DXPS in magenta, different residues between structures are labeled using the same color coding. **B** *pa*DXPS in cyan overlaid with *kp*DXPS in grey and *ec*DXPS in green. The difference in residues between structures is labeled using the same color coding. **C** *pa*DXPS in cyan overlaid with *kp*DXPS in grey and *mt*DXPS in orange, differences in residues between structures are labeled using the same color coding. Residue atoms are shown as sticks (C, cyan; O, red; N, blue; P, orange; S, yellow). Hydrogen bonds are shown as yellow dashed lines, and distances are calculated in Ångstrom.

LC-MS spectroscopy:

LC-MS measurements to determine the intact protein mass were performed using a Dionex Ultimate 3000 RSLC system equipped with an Aeris Widepore XB-c8 (150 x 2.1 mm, 3.6 μm particle diameter(dp)) column (Penomenex, USA). The LC was coupled to a maXis 4G high-resolution time-of-flight (HR-ToF) mass spectrometer (Bruker Daltonic, Germany) using an Apollo electrospray ionization (ESI) source. Separation of 5 μL protein sample on the LC was achieved by a linear gradient from solvent A (H₂O plus 0.1% formic acid) to solvent B (acetonitrile plus 0.1% formic acid) as follows: 0 – 0.5 min (2%), 0.5 – 10.0 min (2 – 75%), 10.0 – 13.0 min (75%), 13.0 – 18.0 min (2%). The following conditions were used for mass spectrometry: capillary voltage 4,000 V, temperature 200 °C, dry gas flow rate 5 L / min and nebulizer 14.5 psi. Data were recorded in a mass range of 150 – 2,500 m/z. Sodium formate clusters were used for calibration of the maXis 4G spectrometer before every injection to avoid mass drift.

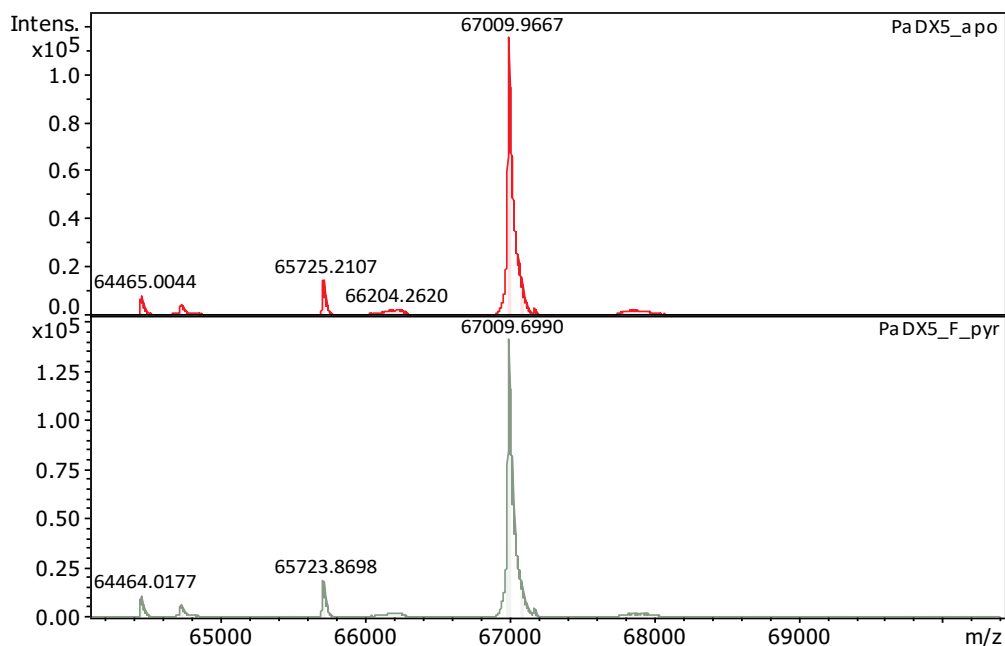


Figure S5: Mass spectra of *paDXPS* before and after incubation with fluoropyruvate. No change in mass between treated and untreated protein in desaturated conditions indicates no covalent modification.

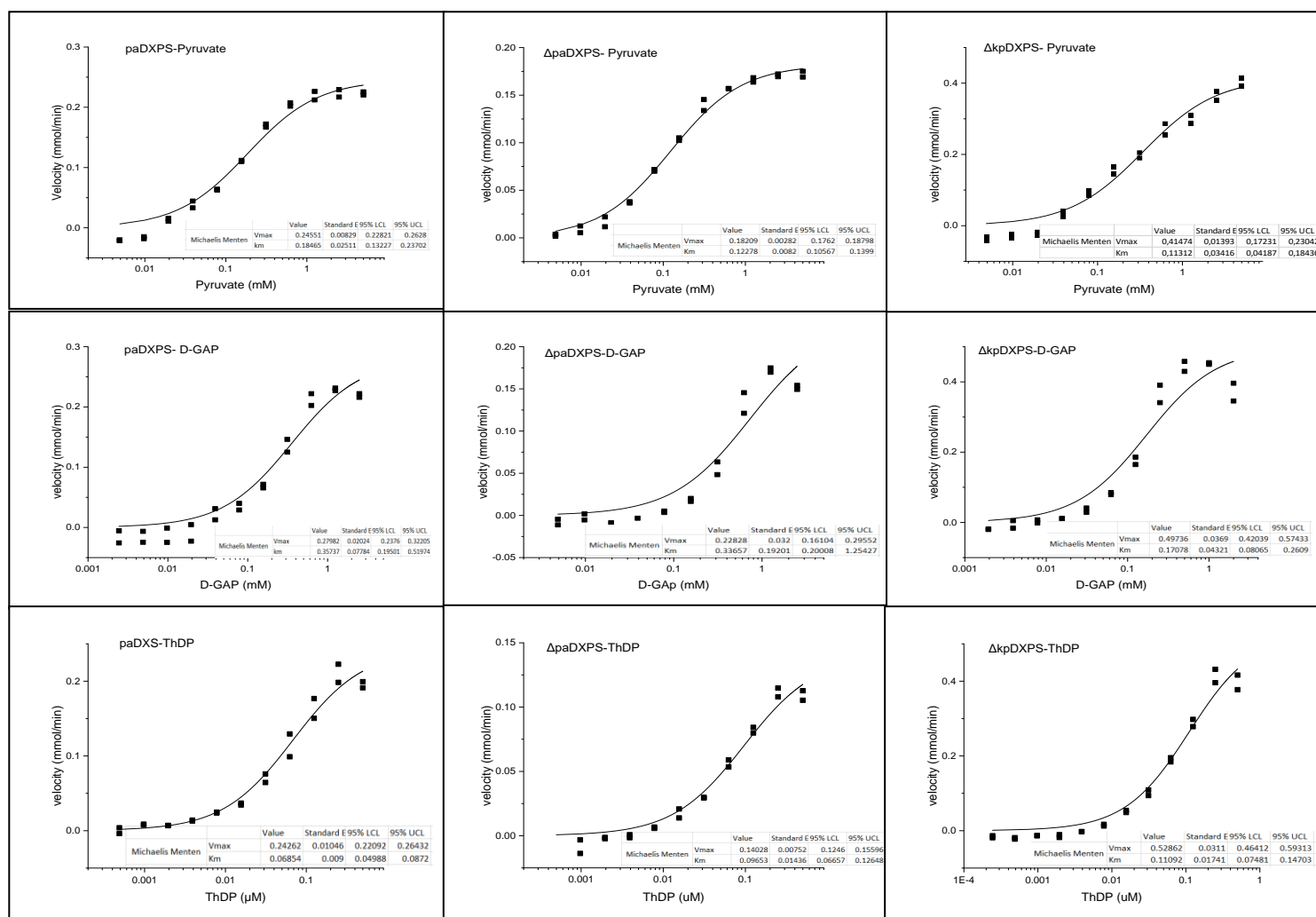


Figure S6: Michaelis-Menten kinetics analysis, K_m values of paDXPS and kpDXPS enzymes, analysis was conducted by varying pyruvate, D-GAP or ThDP concentration, enzyme concentration used: 150nM and 200 nM respectively. When a substrate or cofactor was kept constant, a concentration of 400nM was used for ThDP and 2 mM for pyruvate and D-GAP

Table S3: MOI of fluoropyruvate: IC₅₀ values calculated with varying concentrations of pyruvate (0.8–10 fold K_m), increasing in IC₅₀ with increasing concentration of pyruvate indicating competitive inhibition. D-GAP and ThDP were kept at 2mM and 400nM in all measurements, respectively.

[Pyruvate] (mM)	Fluoropyruvate IC ₅₀ (μ M)
4	154.9 \pm 8.2
2	77.4 \pm 9.8
1	42.6 \pm 1.0
0.5	30.3 \pm 1.3
0.25	27.0 \pm 2.8
0.125	25.9 \pm 1.3
0.0625	22.9 \pm 3.2

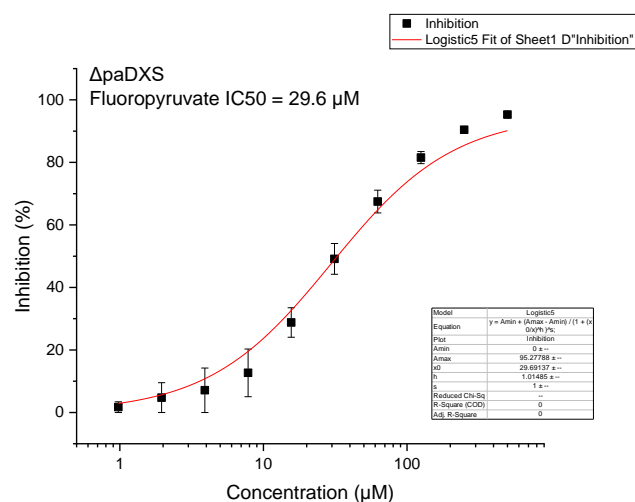
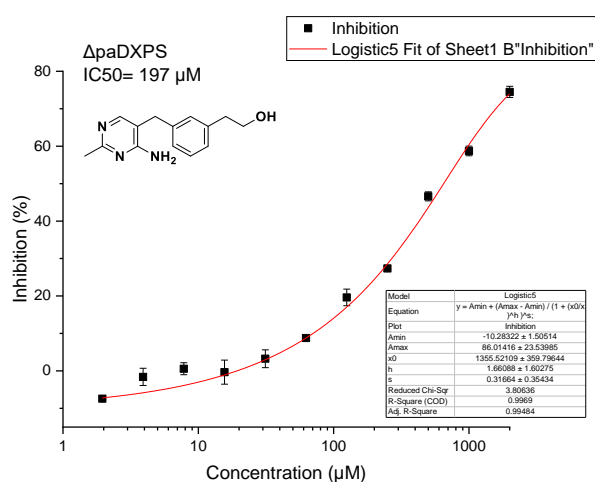
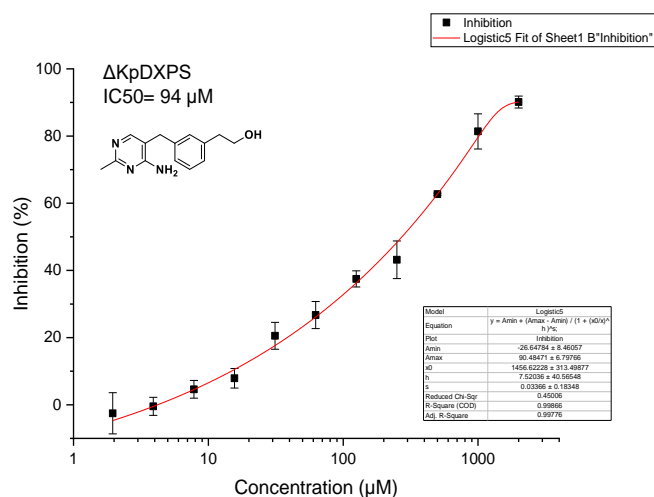


Figure S7: Inhibitory dose-response curves to determine the IC₅₀ for each inhibitor on *pa*DXPS and *kp*DXPS. The curves were fit using non-linear regression analysis program in OriginPro. In the case of the thiamine analog inhibitor ThDP concentration was kept at 4X_{km} (400nM) to insure the enzyme reaches maximum velocity while allowing competitive inhibition. In the case of fluoropyruvate, pyruvate concentration was also kept at 4X _{km} (500μM) for the same reason, 2mM D-Gap was used in both cases.

Chapter 2.2

Design of thiamine analogues for inhibition of thiamine diphosphate (ThDP)-dependent enzymes: Systematic investigation through Scaffold-Hopping and C2-Functionalisation



Design of thiamine analogues for inhibition of thiamine diphosphate (ThDP)-dependent enzymes: Systematic investigation through Scaffold-Hopping and C2-Functionalisation

Alex H.Y. Chan^a, Terence C.S. Ho^a, Rimsha Irfan^a, Rawia A.A. Hamid^{b,c}, Emma S. Rudge^a, Amjid Iqbal^{a,d}, Alex Turner^a, Anna K.H. Hirsch^{b,c}, Finian J. Leeper^{a,*}

^a Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

^b Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123 Saarbrücken, Germany

^c Department of Pharmacy, Saarland University, Campus Building E8.1, 66123 Saarbrücken, Germany

^d Department of Biochemistry, Science Unit, Deanship of Educational Services, Qassim University, Saudi Arabia

ARTICLE INFO

Keywords:

Thiamine diphosphate
Enzyme
Inhibition
Crystal structure
1-deoxy-D-xylulose 5-phosphate synthase

ABSTRACT

Thiamine diphosphate (ThDP), the bioactive form of vitamin B₁, is an essential coenzyme needed for processes of cellular metabolism in all organisms. ThDP-dependent enzymes all require ThDP as a coenzyme for catalytic activity, although individual enzymes vary significantly in substrate preferences and biochemical reactions. A popular way to study the role of these enzymes through chemical inhibition is to use thiamine/ThDP analogues, which typically feature a neutral aromatic ring in place of the positively charged thiazolium ring of ThDP. While ThDP analogues have aided work in understanding the structural and mechanistic aspects of the enzyme family, at least two key questions regarding the ligand design strategy remain unresolved: 1) which is the best aromatic ring? and 2) how can we achieve selectivity towards a given ThDP-dependent enzyme? In this work, we synthesise derivatives of these analogues covering all central aromatic rings used in the past decade and make a head-to-head comparison of all the compounds as inhibitors of several ThDP-dependent enzymes. Thus, we establish the relationship between the nature of the central ring and the inhibitory profile of these ThDP-competitive enzyme inhibitors. We also demonstrate that introducing a C2-substituent onto the central ring to explore the unique substrate-binding pocket can further improve both potency and selectivity.

1. Introduction

Thiamine (I, vitamin B₁) is a water-soluble vitamin essential for energy metabolism. Thiamine deficiency leads to disorders such as beriberi and Wernicke-Korsakoff syndrome [1]. In the majority of mammalian tissues, thiamine diphosphate (II, ThDP) is the most abundant thiamine derivative, accounting for approximately 80% of total thiamine (Fig. 1a) [2–3]. The remainder consists of free thiamine (I), thiamine monophosphate (III, ThMP), thiamine triphosphate (IV, ThTP), adenosine thiamine diphosphate and adenosine thiamine triphosphate. ThDP (the bioactive form of I) is a coenzyme of ThDP-dependent enzymes [4], an extensive family of enzymes that includes pyruvate dehydrogenase complex E1-subunit [5,6], (PDH E1), pyruvate decarboxylase [6] (PDC), pyruvate oxidase [7] (PO), 1-deoxy-D-xylulose

5-phosphate synthase [7–13] (DXPS) and 2-oxoglutarate dehydrogenase complex E1-subunit [14,15], (OGDH E1) (Fig. 1b). Despite the different reactions catalysed, these enzymes all share a similar general mechanism, outlined in Fig. 1a, in which a catalytically active ThDP-ylide V (formed by deprotonation at C2 of the thiazolium ring) chemically converts the substrate and regenerates itself upon product release in an iterative cycle [4,16–17]. Unlike the ThDP-binding pocket, which is structurally similar in all ThDP-dependent enzymes [4], the substrate pocket where substrates bind and chemistry takes place is different for the various enzymes (Fig. 1).

A common strategy in studying ThDP-dependent enzymes lies in the use of thiamine/ThDP analogues for chemical inhibition [18–22] (a pharmacophore is outlined in Fig. 1c). Such analogues are ThDP-competitive inhibitors and generally feature a neutral central aromatic

* Corresponding author.

E-mail address: fjl1@cam.ac.uk (F.J. Leeper).

<https://doi.org/10.1016/j.bioorg.2023.106602>

Received 17 March 2023; Received in revised form 3 May 2023; Accepted 5 May 2023

Available online 10 May 2023

0045-2068/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

ring to capture the strong stabilising interactions between the enzyme's hydrophobic region and the catalytically active high-energy ThDP-ylide **V** [23–30]. There are also reported thiamine/ThDP analogues having the central aromatic ring replaced with a saturated heterocycle [31] or with a planar non-aromatic motif [32–36] but the principle is still to capture the hydrophobic interactions. In the past decade, over 300 such analogues as ligands of thiamine/ThDP-dependent enzymes and/or proteins have been reported in 27 research articles, and even more in patents [29–30,32,36–59]. Despite differing significantly in the substructures on the pyrimidine (R) and the hydroxyethyl tail (R'), the derivatives feature a limited number of central rings. The relative inhibitory activities of the different central rings is difficult to ascertain due to the different assay conditions used and different groups attached. Triazole-based thiamine/ThDP analogues (Fig. 1c), first reported by us in 2008 [27], are the most represented scaffold, due to the ease of constructing the triazole through click chemistry. In our recent paper, several matched molecular pairs revealed that changing the central ring from triazole to furan consistently improved binding affinities [58]. This prompted us to look into the role of the central ring of thiamine/ThDP analogues in inhibiting ThDP-dependent enzymes.

In the current study, we develop ThMP analogues as ThDP-competitive enzyme inhibitors. Using these, we investigate the link

between the central ring and the level of enzyme inhibition by evaluating the inhibition by all the aromatic central rings of thiamine/ThMP/ThDP analogues, reported in the literature since 2013, on multiple ThDP-dependent enzymes. We also demonstrate that introduction of C2-substituents onto the central ring can improve affinity and selectivity towards particular ThDP-dependent enzymes by probing the substrate-binding pocket. Our findings provide valuable structure–activity relationship (SAR) information to guide chemical biologists and medicinal chemists to develop inhibitors for ThDP-dependent enzymes.

2. Results and discussion

2.1. Identity of the central ring affects binding affinity towards ThDP-dependent enzymes

Derivatives of five central scaffolds have been reported in the past decade: thiophene (a) [25,40,60], furan (b) [29,58], triazole (c) [27], pyrrole (d) [30] and *meta*-disubstituted benzene (e) [26,40] (Scheme 1). In this work, we synthesised anew the thiamine analogues of these scaffolds (1a–e) according to Schemes S1–5, improving the yields in some cases. We also prepared a new *para*-disubstituted benzene-based thiamine analogue **1f** (Scheme S6) to explore what effect the angle

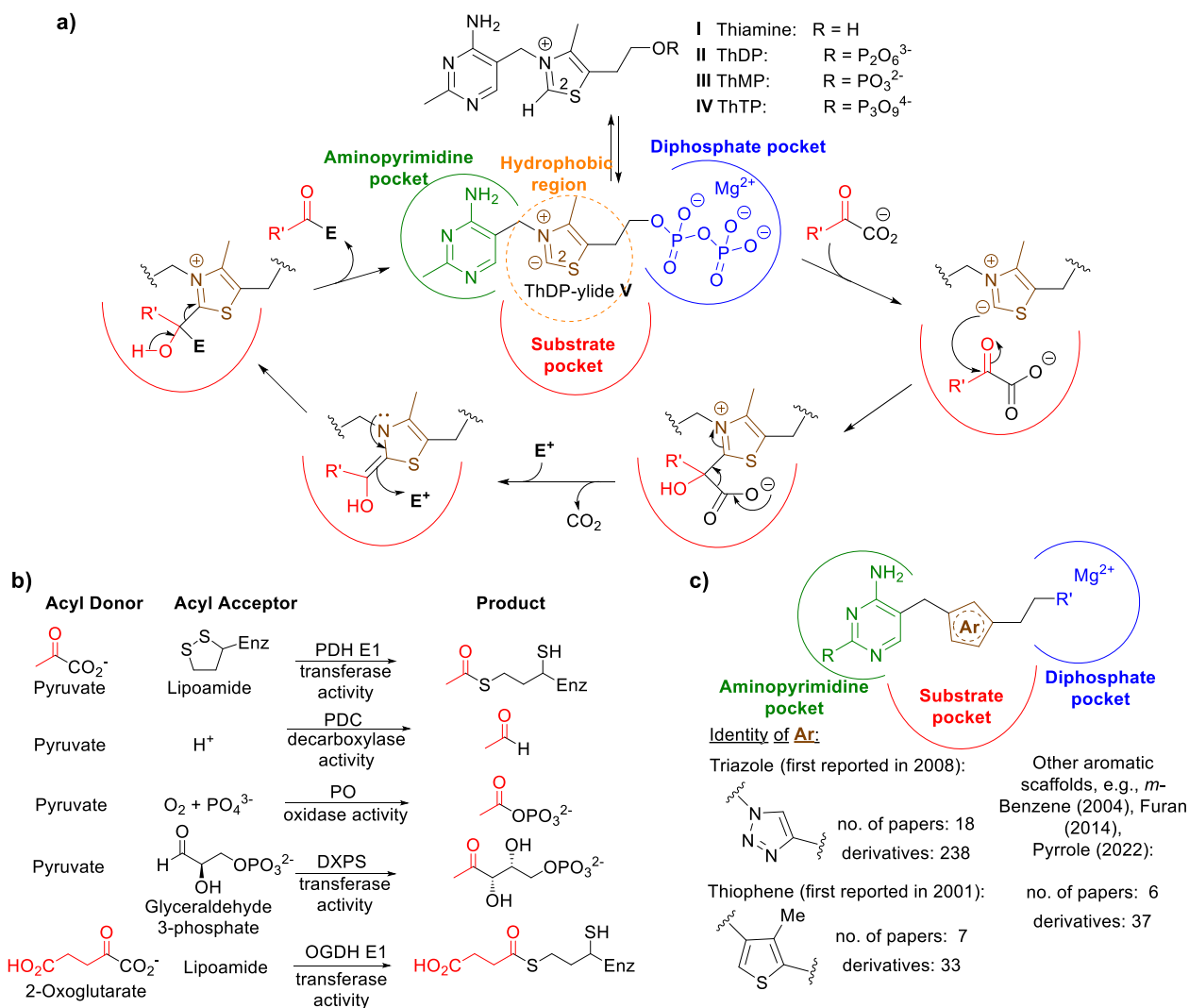
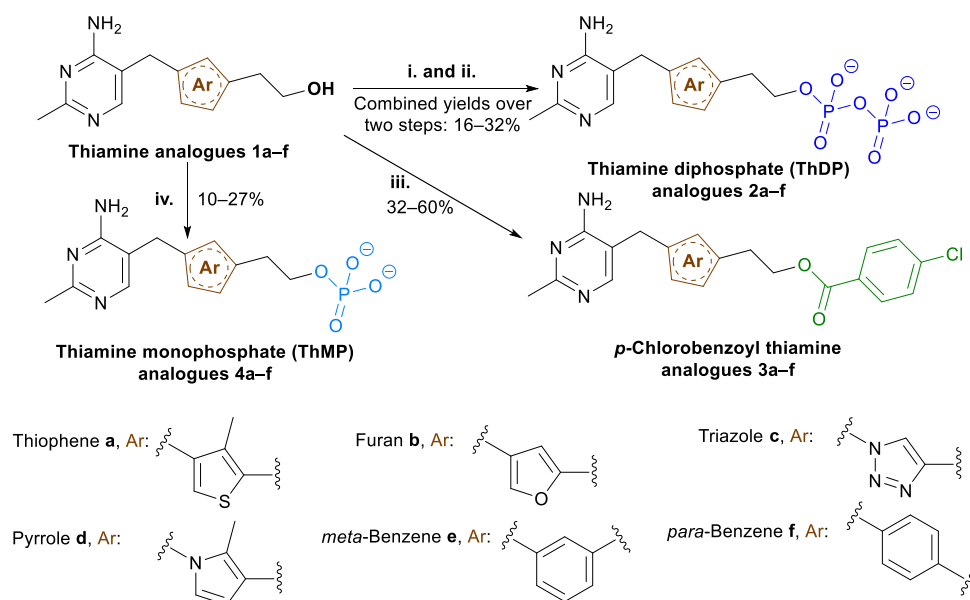


Fig. 1. (a) Schematic depiction of the active site and general mechanism of thiamine diphosphate (ThDP)-dependent enzymes. [15] (b) Reactions catalysed by the enzymes studied in this work. (c) Pharmacophore of thiamine/ThDP analogues and number of papers and analogues of each type reported in the last decade (patents are excluded).



Scheme 1. Tail modifications of thiamine analogues. Reagents and conditions: i) *p*-TslCl, pyridine, 0 °C or RT; ii) (Bu₄N)₃HP₂O₇, MeCN, 0 °C or RT; iii) *p*-chlorobenzoic acid, DCC, DMAP, DMF, RT; iv) POCl₃, RT. Thiamine analogues **1a–f** were obtained as shown in Schemes S1–6.

Table 1
Summary of inhibitory activities (K_i) of compounds 1–4.^a

Compound		K_i			
Tail	Ar	PDH E1	PDC	PO	OGDH E1
 1a–f	Thiophene a	18 nM	63 μM	24 μM	7.6 μM
	Furan b	110 nM	78 μM	83 μM	10 μM ⁵⁸
	Triazole c	150 nM	190 μM	170 μM	ND
	Pyrrole d	96 nM	72 μM	20 μM	ND
	<i>m</i> -Benzene e	25 nM	67 μM	38 μM	ND
	<i>p</i> -Benzene f	34 nM	70 μM	45 μM	ND
 2a–f	Thiophene a	33 pM	100 nM	1.7 nM	5–6 nM ²⁶
	Furan b	59 pM	570 nM	92 nM	78 nM
	Triazole c	67 pM	680 nM	150 nM	95 nM
	Pyrrole d	NA	NA	NA	NA
	<i>m</i> -Benzene e	37 pM	140 nM	7.8 nM	ND
	<i>p</i> -Benzene f	46 pM	220 nM	19 nM	ND
 3a–f	Thiophene a	8.0 nM	3a–f are selective to PDH E1.		
	Furan b	20 nM	When tested under assay conditions with [ThDP]: [compound] = 1:5, percentage inhibition was < 10%		
	Triazole c	60 nM			
	Pyrrole d	17 nM			
	<i>m</i> -Benzene e	11 nM			
	<i>p</i> -Benzene f	14 nM			
 4a–f	Thiophene a	3.8 nM	44 μM	2.5 μM	2.9 μM
	Furan b	19 nM	60 μM	18 μM	8.3 μM
	Triazole c	33 nM	130 μM	29 μM	17 μM
	Pyrrole d	NA	NA	NA	NA
	<i>m</i> -Benzene e	4.5 nM	33 μM	1.8 μM	4.6 μM
	<i>p</i> -Benzene f	6.6 nM	51 μM	10 μM	5.8 μM

^a Data are the means of measurements in three technical replicates. Since all compounds are thiamine diphosphate (ThDP)-competitive enzyme inhibitors, their K_i values were calculated using to the enzyme's K_M for ThDP: 0.05 μM (PDH E1), 15 μM (yeast PDC), 5 μM (PO) and 3 μM (OGDH E1). Refer to Figures S1–4 for details (including margins of error and assay conditions). Affinities of compounds from previously reported studies were re-determined in the current study for consistency, except where indicated. NA, not available because the syntheses of **2d** and **4d** were unsuccessful. ND, not determined.

between the two substituents has on the inhibition of the enzymes.

Modification (phosphorylation or esterification) of the hydroxyethyl tail of thiamine analogues **1a–f** generally leads to stronger binders towards ThDP-dependent enzymes (Scheme 1). Thiamine and its analogues are weak binders, as the tight binding of ThDP to the enzymes is largely through the ionic interactions with the Mg²⁺ cation in the diphosphate pocket (Fig. 1a). Thiamine analogues **1a–f** were tosylated and then di-phosphorylated to yield ThDP analogues **2**, coupled with *p*-chlorobenzoic acid to yield esters **3a–f** (as we have recently identified **3b** as a potent and selective PDH E1 inhibitor) [58], and

monophosphorylated to yield ThMP analogues **4** (the first reported ThMP analogues). The syntheses of pyrro-ThDP (**2d**) and pyrro-ThMP (**4d**) were unsuccessful, presumably due to the electron-rich pyrrole ring making these compounds (and the intermediates) unstable.

IC₅₀ values of the 22 derivatives were determined for inhibition of four ThDP-dependent enzymes, namely PDH E1, PDC, PO and OGDH E1 (Table 1 and Figures S1–4). Analogues 1–4 were validated as ThDP-competitive inhibitors because their apparent inhibition decreased with increasing levels of ThDP (see SI and previous reports) [26–27,29–30,58]; therefore, their K_i values were calculated based on

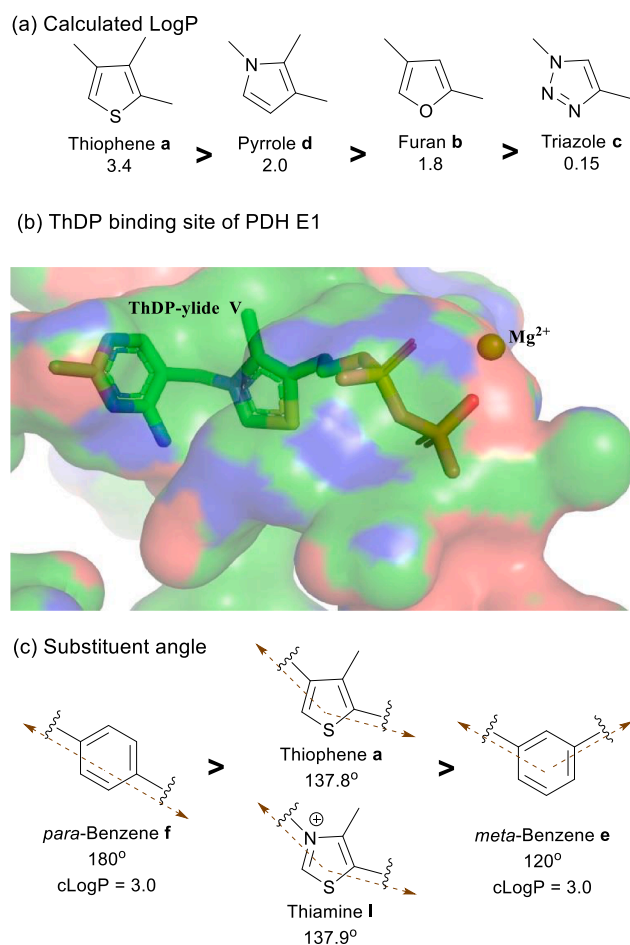
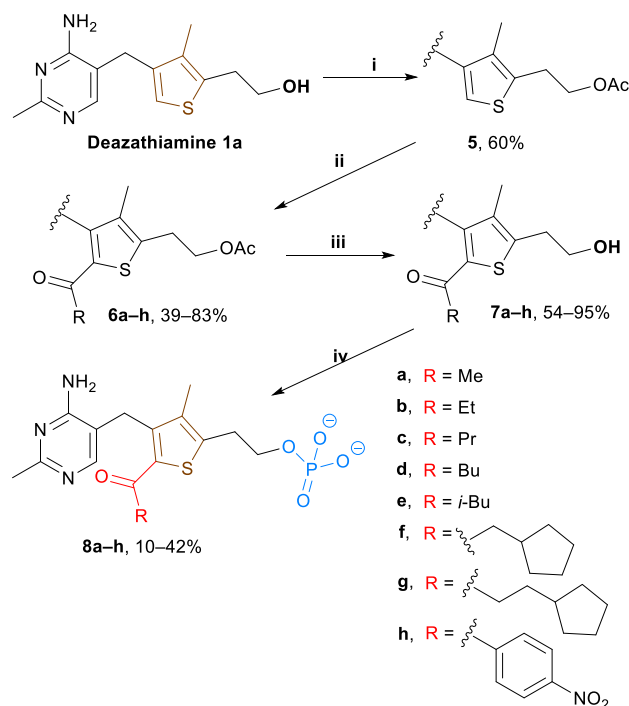


Fig. 2. Suggested explanation of the data in Table 1. (a) Hydrophobicity of the heterocyclic rings. The more hydrophobic the ring the better the binding (cLogP from MarvinSketch 21.1). (b) Crystal structure (PDB 1NI4) of ThDP-bound to human PDH E1; ThDP-binding pocket shown as surface. The thiazolium ring of ThDP is anchored in a central hydrophobic region (green surface) between the polar aminopyrimidine (left) and diphosphate (right) pockets. (c) Comparison of the substituent angles of the thiophene and benzenes. Binding affinity is higher if the angle matches that of thiamine (see Figure S5).

the measured K_M of ThDP for each individual enzyme (PDH E1, 0.05 μM ; yeast PDC, 15 μM ; PO, 5 μM ; OGDH E1, 3 μM) and the [ThDP] used in each assay ($\text{IC}_{50}/[\text{ThDP}] = K_I/K_M(\text{ThDP})$).

The affinity of the analogues increases with the negative charges of the tail of thiamine/ThMP/ThDP analogues **1**, **2** and **4** due to the electrostatic interactions with the Mg^{2+} in the diphosphate pocket, consistent with our previous findings [27]. ThMP analogues **4** bind more strongly than thiamine analogues **1** but much more weakly than ThDP analogues **2**. This trend is consistent for all enzymes and with all scaffolds **a–f** (Table 1). The K_I values of the ThMP analogues **4** (typically in the low-micromolar range) are comparable to the affinities of ThDP, suggesting that the gain in binding energy from the hydrophobic interactions of the neutral central ring compensates for the weakened interactions with the Mg^{2+} cation. As expected, esters **3a–f** are all selective for PDH E1 over the other enzymes, probably because only PDH E1 has a large enough diphosphate pocket to accommodate the ester tail [58].

Affinity of the analogues to all the enzymes tested increases with the hydrophobicity of the central ring due to hydrophobic interactions with the enzyme. The trend of affinity among the five-membered rings is in all but one case: thiophene **a** > pyrrole **d** > furan **b** > triazole **c**. The one exception is **1d**, which has a K_I marginally (though not significantly)



Scheme 2. Synthesis of acyl deaza-ThMP derivatives. Reagents and conditions: i) acetic acid, DCC, DMAP, DMF, RT; ii) acyl chloride, AlCl_3 , DCM, RT; iii) K_2CO_3 , MeOH, RT; iv) POCl_3 , RT.

lower than **1a** for PO. The central ring of analogues **1–4** occupies a hydrophobic region (green in Fig. 2b) in the ThDP pocket that forms strong hydrophobic interactions with the catalytically active high-energy ThDP-ylide V. The dielectric constant of this hydrophobic region has been estimated to be 13–15 [61], which explains why the more hydrophobic rings (as indicated by the cLogP values of the methyl-substituted rings shown in Fig. 2a) bind better.

Binding affinity also seems to be, to some extent, a function of the geometry of the disubstituted central ring. Table 1 shows a general trend in affinity: thiophene **a** > *meta*-benzene **e** > *para*-benzene **f**. As the hydrophobicities of the thiophene and the two benzenes are similar, we attribute the higher affinity of the thiophene to the smaller deviation of its angle between the two substituents from the corresponding angle for the thiazolium ring of thiamine, which is 137.9° (Figure S5). Exceptions to this trend are found for the monophosphates **4a** and **4e**, with **4e** inhibiting PDC and PO slightly more than **4a**. Molecular docking (Figure S6) suggested that the different substituent angle for **4e** allowed the phosphate to be a bidentate ligand for the Mg^{2+} for PDC and PO, whereas for **4a** only monodentate coordination was possible. Of the two factors, hydrophobicity and geometry of the central ring, the former seems to play the larger role: the more hydrophobic *para*-benzene analogues **f** bind more strongly in all cases than any of the equivalent more polar furan **b** or triazole **c** analogues.

2.2. Functionalisation of the central ring improves selectivity towards particular members of ThDP-dependent enzymes

All ThDP-dependent enzymes require the coenzyme ThDP for activity so using ThDP-competitive analogues that are close in structure to ThDP is not expected to lead to selectivity towards individual types of enzymes. However, individual types of enzyme differ significantly in terms of their substrates and catalytic action (Fig. 1b). Targeting the substrate-binding pocket might, therefore, result in selective inhibition; this strategy is recently exemplified by the use of acylphosphonates as DXPS inhibitors showing greater than 15,000-fold selectivity over other

Table 2

 K_I values of compounds **4a** and **8a–h** and their affinities relative to that of ThDP.^a

 8a–h	PDH E1		yeast PDC		bacterial PDC		PO		DXPS	
	K_I	vs ThDP	K_I	vs ThDP	K_I	vs ThDP	K_I	vs ThDP	K_I	vs ThDP
Deaza-ThMP 4a	3.8 nM	13	44 μ M	0.34	0.7 μ M	0.50	2.5 μ M	2.0	1.1 μ M	0.046
8a R = Me	3.2 nM	16	42 μ M	0.36	0.5 μ M	0.68	1.3 μ M	3.8	80 nM	0.63
8b R = Et	2.1 nM	24	Insoluble		0.6 μ M	0.59	3.4 μ M	1.5	60 nM	0.83
8c R = Pr	1.1 nM	46	Insoluble		1.8 μ M	0.19	28 μ M	0.18	310 nM	0.16
8d R = Bu	1.3 nM	39	Insoluble		>3.5 μ M	<0.1	33 μ M	0.15	370 nM	0.14
8e R = <i>i</i> Bu	1.9 nM	27	Insoluble		>3.5 μ M	<0.1	46 μ M	0.11	90 nM	0.56
8f R = CH ₂ cPent	2.0 nM	25	Insoluble		>3.5 μ M	<0.1	>50 μ M	<0.1	100 nM	0.50
8g R = CH ₂ CH ₂ cPent	2.3 nM	22	Insoluble		>3.5 μ M	<0.1	>50 μ M	<0.1	4.0 μ M	0.012
8h R = <i>p</i> -NO ₂ Ph	6.0 nM	8	Insoluble		>3.5 μ M	<0.1	>50 μ M	<0.1	330 nM	0.15

^a Data are the means of measurements in three technical replicates for all enzymes, except for DXPS in two technical replicates. IC₅₀ values were determined at [ThDP] = 10 μ M (PDH E1), 150 μ M (yeast PDC), 10 μ M (bacterial PDC), 60 μ M (PO) and 0.5 μ M (DXPS) and affinity relative to that of ThDP is given by [ThDP]/IC₅₀. K_I is calculated using [ThDP]/IC₅₀ = $K_{M(\text{ThDP})}/K_I$. $K_{M(\text{ThDP})}$ is 350 nM for bacterial PDC and 50 nM for DXPS. Refer to Figures S7–11 for details (including margins of error and assay conditions).

ThDP-dependent enzymes [62]. We therefore aimed to introduce substituents onto the position equivalent to C2 of ThDP (which is where the substrates bind) of our ThMP analogue **4a**. We synthesised eight acyl deaza-ThMP derivatives **8a–h** through Friedel–Crafts acylation reactions on acetate ester **5**, followed by ester hydrolysis and phosphorylation (Scheme 2). We chose to functionalise deaza-ThMP **4a** because thiophene **a** is the most potent among the five-membered rings and it only has one free position on the thiophene, so acylation cannot lead to a regioisomeric mixture. The acyl groups were chosen to explore the length of the substrate-binding pocket (linear alkyl chains **8a–d**) and its breadth (by including branched and cyclic chains **8e–h**).

Table 2 (a summary of the data in Figures S7–11) shows the K_I values of derivatives **8a–h** for PDH E1, PDC, PO and DXPS. Hydrophobic compounds **8b–h** were not sufficiently soluble for the assay conditions of yeast PDC, so a bacterial PDC (from *Zymomonas mobilis*) was also used, due to the better assay conditions (lower [ThDP] needed to saturate the enzyme in the absence of inhibitors) that avoid solubility problems. As the inhibitors are competitive with ThDP, the degree of inhibition is not determined by the [inhibitor] and the K_I value alone but the [ThDP] and the K_M value of ThDP must also be taken into account. Accordingly, the affinity of the inhibitor relative to that of ThDP (= [ThDP]/IC₅₀ = $K_{M(\text{ThDP})}/K_I$) is important and is also shown in Table 2. To assess the selectivity of an inhibitor for one enzyme over another, it is these relative affinity values that should be compared as the [ThDP] is the same for both enzymes. Notably, pyruvate is the acyl donor substrate for all these enzymes, so part of the substrate pocket is likely to be similar among the enzymes, making selective inhibition more challenging, but the acyl acceptor substrates differ markedly. The data in Table 2 show that, even without the C2 acyl group, these ThMP analogues are selective for PDH E1 over the other enzymes – for **4a** the selectivities for PDH E1 are 26, 6.5 and 283-fold over bacterial PDC, PO and DXPS respectively. On all tested enzymes, a two-carbon acetyl group (**8a**) enhances affinity by mimicking the ThDP-pyruvate intermediate after decarboxylation (Fig. 1a). However, a branched, bulky group (as in **8e**) favours binding to PDH E1 and DXPS but virtually abolishes the activity on PDC and PO, with the selectivity for PDH E1 being over 200-fold for both. The natural acyl acceptor substrates of PDC and PO are much smaller than for PDH E1, so the enhanced selectivity is probably because the C2 acyl group is too large to be accommodated in their substrate pocket. For DXPS the best inhibitor is **8b**, with the three carbon propionyl group: for this analogue the selectivity for PDH E1 over DXPS has been reduced to only 25-fold.

The key finding from Table 2 is that using C2-functionalised analogues to occupy the ThDP pocket and probe the substrate pocket

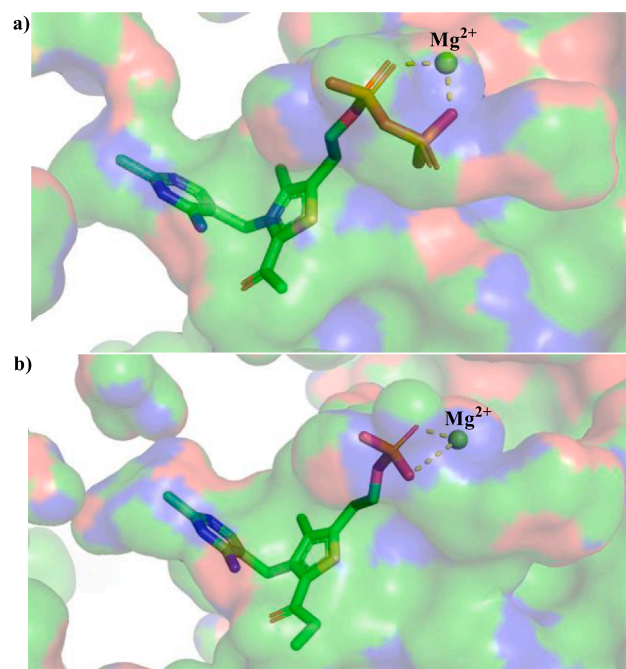


Fig. 3. (a) Reported structure of 2-acetyl ThDP bound to *Deinococcus radiodurans* DXPS (PDB: 6OUW). (b) Crystal structure of **8b** bound to *Klebsiella pneumoniae* DXPS obtained in this work (PDB: 8BZX). ThDP-binding pocket shown as surface.

simultaneously is one way to enhance selectivity for particular ThDP-dependent enzymes. However, to achieve target-specificity, C2-functionalisation alone may not be enough. In recent research, we applied C2-functionalisation to ester **3b** and obtained several derivatives that showed nanomolar activities towards PDH E1 but were inactive on other ThDP-dependent enzymes [58]; this type of inhibitor has potential as a cellular probe and/or treatment for cancer cell lines that overexpress PDH [63–64].

As DXPS is essential to many pathogenic microorganisms, but is absent in mammals, there has been considerable interest in addressing DXPS as an antimicrobial target in recent years [11–13]. To gain structural insights to aid ligand development, we recently solved the first crystal structures of DXPS from *Klebsiella pneumoniae* [65]. In this

work, we obtained an X-ray crystal structure of the binary complex between the KpDXPS and **8b**. As a reference, Fig. 3a shows a reported structure [11] of a DXPS holoenzyme (from *Deinococcus radiodurans*) in which a 2-acetyl ThDP coenzyme is bound: the coenzyme shows the characteristic V-shaped conformation (between the aminopyridine and the thiazolium ring) and the acetyl group defines the position of the substrate pocket. Fig. 3b shows that ThMP analogue **8b** binds in the ThDP pocket with the same characteristic V-shaped conformation, confirming that it is ThDP-competitive. Furthermore, the 5-acyl group occupies the position of the acetyl group, thus supporting our expected pharmacophore where this C2-group extends into the substrate pocket (Fig. 1c). The monophosphate **8b** interacts with the Mg²⁺ cation with two of the oxygen atoms of the phosphate (bond distances 1.9 and 2.9 Å), whereas 2-acetyl ThDP interacts with one oxygen atom from each phosphoryl group (bond distances 2.0 and 2.1 Å). Selected compounds were computationally docked into the active sites of PDH E1, PDC and PO (Figure S12), and all these structures showed very similar binding modes to that found in this crystal structure.

We have here introduced ThMP analogues **4** as ThDP-competitive inhibitors, and used them to investigate the effects of switching the central scaffold (**4a–f** in Table 1) and C2-functionalisation (**8a–h** in Table 2) on affinity and to obtain a crystal structure. We envision that they may have potential to be applied in cell-based systems because, firstly, there are no (identified) biological targets for thiamine monophosphate in humans (other than transporter SLC19A1 for ThMP and ThDP efflux), which reduces the likelihood of ThMP analogues causing off-target effects. Secondly, ThMP analogues, though probably not very cell-permeable, are more amenable to being administered as prodrugs than ThDP analogues. There are many clinically approved drugs that are prodrugs masking monophosphates, whereas development of prodrugs for diphosphates is slow [66].

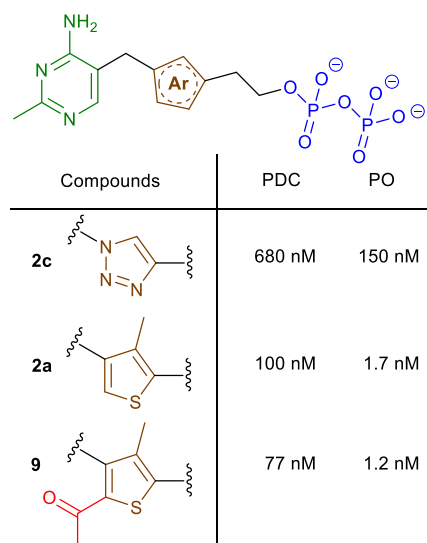
2.3. Scaffold-hopping and C2-functionalisation of ThDP analogues: From sub-micromolar to low-nanomolar affinities

We were interested in discovering whether adding a C2-acyl group could improve the affinity of the ThDP analogues such as deazaThDP **2a** (as well as the ThMP analogues described above). The C2-acetyl derivative of deazaThDP **9** was synthesised by a second phosphorylation of deazaThMP **8a** (Scheme S7). The K_i values of **9** for PDC and PO are shown in Fig. 4 (summarising data from Figure S13). Triazole-based ThDP analogue **2c** is a strong inhibitor, but deaza-ThDP **2a** and its C2-acetylated derivative **9** are much more potent inhibitors. C2-acetylated **9** is the most potent of the ThDP analogues and binds about 9 times and 125 times more strongly than **2c** to PDC and PO, respectively. Although the triazole scaffold is easy to make, the data in Tables 1–2 and Fig. 4 show that switching to a more hydrophobic scaffold that bears an accessible C2-position and introducing a suitable C2-substituent can enormously improve both affinity (due to the stronger hydrophobic interactions) and selectivity (due to the different steric effects).

3. Conclusions

Thiamine/ThDP analogues have long been established as ThDP-competitive inhibitors of ThDP-dependent enzymes [26–27,29–30,58]; in this work, we report novel thiamine monophosphate analogues as ThDP-competitive inhibitors with affinities similar to those of ThDP. We also report a crystal structure of an enzyme (DXPS) with one of one of these ThMP analogues bound. The lack of (known) intracellular ThMP binding targets [21] and the synthetically accessible prodrugs for monophosphates collectively increase the applicability of ThMP analogues in cell-based systems to study/manipulate ThDP-dependent enzymes and their pathways.

Using the ThMP analogues, we show that the triazole scaffold, which is the most represented in the literature, has the lowest affinity amongst



Compounds	PDC	PO
2c	680 nM	150 nM
2a	100 nM	1.7 nM
9	77 nM	1.2 nM

Fig. 4. Summary of inhibition constants (K_i) of compounds **2a, **2c** and **9**.** Data are the means of measurements in three technical replicates. Since all compounds are ThDP-competitive enzyme inhibitors, their inhibition constants (K_i) are calculated from the measured IC_{50} and the K_M for ThDP of 15 μ M (yeast PDC) and 5 μ M (PO). Refer to Figure S13 for details (including margins of error and assay conditions).

the six aromatic scaffolds. We propose that the affinity of the central ring is mainly a function of the hydrophobicity and secondarily of the substituent geometry. We also demonstrate that adding a C2-substituent to explore the unique substrate pocket, based on the substrates of a particular ThDP-dependent enzyme, can improve not only affinity, but also selectivity, even between enzymes that utilise the same donor substrate. While these concepts have been exemplified with five enzymes, they could in theory be extended to other ThDP-dependent enzymes because thiamine/ThMP/ThDP analogues all share the same principle of enzyme inhibition.

We anticipate that our findings will convince researchers that although derivatives based on the triazole scaffold are easy to prepare, switching to another more hydrophobic scaffold that bears an available C2-position may lead to compounds with an improved inhibition profile. Having said that, the triazole and the thiophene scaffolds (and others as well) are not in competition, but are complementary: the easy-to-make triazole-based analogues can be used to get preliminary insights into the target ThDP-dependent enzyme (e.g. its polarity, thus aqueous solubility, makes obtaining a crystal structure easier), and then scaffold-hopping and C2-functionalisation can be explored in the lead optimisation process. We believe that the insights into the relationship between the structure and the inhibitory profile will aid development of better analogues, ultimately expanding our toolbox for studying ThDP-dependent enzymes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful for funding from K. M. Medhealth (AHYC and TCSH), the University of Cambridge Harding Distinguished Postgraduate Scholars Programme Leverage Scheme (RI), the Cambridge Commonwealth Trust (AD), the Schlumberger Foundation Faculty for the Future Fellowship (RAAH) and the Helmholtz Association Initiative and Networking Fund and the European Research Council, ERC Starting grant 757913 (AKHH).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bioorg.2023.106602>.

References

- E. Rees, L.R. Gowing, Supplementary thiamine is still important in alcohol dependence, *Alcohol Alcohol.* 48 (2013) 88–92, <https://doi.org/10.1093/alcal/ags120>.
- L. Bettendorff, B. Lakaye, G. Kohn, P. Wins, Thiamine triphosphate: a ubiquitous molecule in search of a physiological role, *Metab. Brain Dis.* 29 (2014) 1069–1082, <https://doi.org/10.1007/s11011-014-9509-4>.
- V.I. Bunik, V.A. Aleshin, Analysis of the Protein Binding Sites for Thiamin and Its Derivatives to Elucidate the Molecular Mechanisms of the Noncoenzyme Action of Thiamin (Vitamin B1), *Stud. Nat. Prod. Chem.* 53 (2017) 375–429, <https://doi.org/10.1016/B978-0-444-63930-1.00011-9>.
- R.A.W. Frank, F.J. Leeper, B.F. Luisi, Structure, mechanism and catalytic duality of thiamine-dependent enzymes, *Cell. Mol. Life Sci.* 64 (2007) 892–905, <https://doi.org/10.1007/s00018-007-6423-5>.
- N.S. Nemeria, P. Arjunan, K. Chandrasekhar, M. Mossad, K. Tittmann, W. Furey, F. Jordan, Communication between Thiamin Cofactors in the *Escherichia coli* Pyruvate Dehydrogenase Complex E1 Component Active Centers: Evidence for a “Direct Pathway” between the 4'-Aminopyrimidine N1' Atoms, *J. Biol. Chem.* 285 (2010) 11197–11209, <https://doi.org/10.1074/jbc.M109.069179>.
- S. Prajapati, D. Haselbach, S. Wittig, M.S. Patel, A. Chari, C. Schmidt, H. Stark, K. Tittmann, Structural and functional analyses of the human PDH complex suggest a “Division-of-Labor” mechanism by local E1 and E3 clusters, *Structure* 27 (2019) 1124–1136, <https://doi.org/10.1016/j.str.2019.04.009>.
- M. Paulikat, C. Wechsler, K. Tittmann, R.A. Mata, Theoretical studies of the electronic absorption spectra of thiamin diphosphate in pyruvate decarboxylase, *Biochemistry* 56 (2017) 1854–1864, <https://doi.org/10.1021/acs.biochem.6b00984>.
- S. Dai, L.-M. Funk, F.R. von Pappenheim, V. Sautner, M. Paulikat, B. Schröder, J. Uranga, R.A. Mata, K. Tittmann, Low-barrier hydrogen bonds in enzyme cooperativity, *Nature* 573 (2019) 609–613, <https://doi.org/10.1038/s41586-019-1581-9>.
- J.K. White, S. Handa, S.L. Vankayala, D.J. Merkler, H.L. Woodcock, Thiamin Diphosphate Activation in 1-Deoxy-D-xylulose 5-Phosphate Synthase: Insights into the Mechanism and Underlying Intermolecular Interactions, *J. Phys. Chem. B* 120 (2016) 9922–9934, <https://doi.org/10.1021/acs.jpcc.6b07248>.
- A.A. DeColli, N.S. Nemeria, A. Majumdar, G.J. Gerfen, F. Jordan, C.L. Freil Meyers, Oxidative decarboxylation of pyruvate by 1-deoxy-D-xylulose 5-phosphate synthase, a central metabolic enzyme in bacteria, *J. Biol. Chem.* 293 (2018) 10857–10869, <https://doi.org/10.1074/jbc.RA118.001980>.
- P.-Y.-T. Chen, A.A. DeColli, C.L. Freil Meyers, C.L. Drennan, X-ray crystallography-based structural elucidation of enzyme-bound intermediates along the 1-deoxy-D-xylulose 5-phosphate synthase reaction coordinate, *J. Biol. Chem.* 294 (2019) 12405–12414, <https://doi.org/10.1074/jbc.RA119.009321>.
- A.A. DeColli, X. Zhang, K.L. Hefflin, F. Jordan, C.L. Freil Meyers, Active Site Histidines Link Conformational Dynamics with Catalysis on Anti-Infective Target 1-Deoxy-D-xylulose 5-Phosphate Synthase, *Biochemistry* 58 (2019) 4970–4982, <https://doi.org/10.1021/acs.biochem.9b00878>.
- M.L. Johnston, E.M. Bonett, A.A. DeColli, C.L. Freil Meyers, Revealing Donor Substrate-Dependent Mechanistic Control on DXPS, an Enzyme in Bacterial Central Metabolism, *Biochemistry* 61 (2022) 1810–1823, <https://doi.org/10.1021/acs.biochem.1c00019>.
- N.S. Nemeria, X. Zhang, J. Leandro, J. Zhou, L. Yang, S.M. Houten, F. Jordan, Toward an understanding of the structural and mechanistic aspects of protein-protein interactions in 2-oxoacid dehydrogenase complexes, *Life* 11 (2021) 407, <https://doi.org/10.3390/life11050407>.
- J. Chakraborty, N. Nemeria, Y. Shim, X. Zhang, E.L. Guevara, H. Patel, E.T. Farinas, F. Jordan, Engineering the 2-oxoglutarate dehydrogenase complex to understand catalysis and alter substrate recognition, *Reactions* 3 (2022) 139–159, <https://doi.org/10.3390/reactions3010011>.
- S. Prajapati, F. Rabe von Pappenheim, K. Tittmann, Frontiers in the enzymology of thiamin diphosphate-dependent enzymes, *Curr. Opin. Struct. Biol.* 76 (2022), 102441, <https://doi.org/10.1016/j.sbi.2022.102441>.
- D. Kern, G. Kern, H. Neef, K. Tittmann, M. Killenberg-Jabs, C. Wikner, G. Schneider, G. Hübner, How thiamine diphosphate is activated in enzymes, *Science* 275 (1997) 67–70, <https://doi.org/10.1126/science.275.5296.67>.
- C.L. Berthold, D. Gocke, M.D. Wood, F.J. Leeper, M. Pohl, G. Schneider, Crystal structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemo- and enantioselective carbonylation reaction, *Acta Crystallogr. D Biol. Crystallogr.* 63 (2007) 1217–1224, <https://doi.org/10.1107/S0907444907050433>.
- C.L. Berthold, C.G. Toyota, P. Moussatche, M.D. Wood, F. Leeper, N.G.J. Richards, Y. Lindqvist, Crystallographic snapshots of oxalyl-CoA decarboxylase give insights into catalysis by nonoxidative ThDP-dependent decarboxylases, *Structure* 15 (2007) 853–861, <https://doi.org/10.1016/j.str.2007.06.001>.
- W. Versées, S. Spaepen, M.D.H. Wood, F.J. Leeper, J. Vanderleyden, J. Steyaert, Molecular Mechanism of Allosteric Substrate Activation in a Thiamine Diphosphate-dependent Decarboxylase, *J. Biol. Chem.* 282 (2007) 35269–35278, <https://doi.org/10.1074/jbc.M706048200>.
- S. Lüdtke, P. Neumann, K.M. Erixon, F. Leeper, R. Kluger, R. Ficner, K. Tittmann, Sub-ångström-resolution crystallography reveals physical distortions that enhance reactivity of a covalent enzymatic intermediate, *Nat. Chem.* 5 (2013) 762–767, <https://doi.org/10.1038/nchem.1728>.
- A. Tylicki, Z. Lotowski, M. Siemieniuk, A. Ratkiewicz, Thiamine and selected thiamine antivitamin — biological activity and methods of synthesis, *Biosci. Rep.* 38 (2018), <https://doi.org/10.1042/BSR20171148>.
- D.S. Shreve, M.P. Holloway, J.C. Haggerty, H.Z. Sable, The catalytic mechanism of transketolase. Thiamin pyrophosphate-derived transition states for transketolase and pyruvate dehydrogenase are not identical, *J. Biol. Chem.* 258 (1983) 12405–12408, [https://doi.org/10.1016/S0021-9258\(17\)44189-5](https://doi.org/10.1016/S0021-9258(17)44189-5).
- N. Nemeria, Y. Yan, Z. Zhang, A.M. Brown, P. Arjunan, W. Furey, J.R. Guest, F. Jordan, Inhibition of the *Escherichia coli* Pyruvate Dehydrogenase Complex E1 Subunit and Its Tyrosine 177 Variants by Thiamin 2-Thiazolone and Thiamin 2-Thiothiazolone Diphosphates: Evidence for Reversible Tight-binding Inhibition, *J. Biol. Chem.* 276 (2001) 45969–45978, <https://doi.org/10.1074/jbc.M104116200>.
- D. Hawksley, D.A. Griffin, F.J. Leeper, Synthesis of 3-deazathiamine, *J. Chem. Soc., Perkin Trans. 1* (2001) 144–148, <https://doi.org/10.1039/B006962K>.
- S. Mann, C. Perez Melero, D. Hawksley, F.J. Leeper, Inhibition of thiamin diphosphate dependent enzymes by 3-deazathiamin diphosphate, *Org. Biomol. Chem.* 2 (2004) 1732, <https://doi.org/10.1039/B403619K>.
- K.M. Erixon, C.L. Dabalos, F.J. Leeper, Synthesis and biological evaluation of pyrophosphate mimics of thiamine pyrophosphate based on a triazole scaffold, *Org. Biomol. Chem.* 6 (2008) 3561, <https://doi.org/10.1039/B806580B>.
- A.A. Thomas, J. De Meese, Y. Le Huerou, S.A. Boyd, T.T. Romoff, S.S. Gonzales, I. Gunawardana, T. Kaplan, F. Sullivan, K. Sondroski, J.P. Lyssikatos, T.D. Aicher, J. Ballard, B. Bernat, W. DeWolf, M. Han, C. Lemieux, D. Smith, S. Weiler, S. K. Wright, G. Vigers, B. Brandhuber, Non-charged thiamine analogs as inhibitors of enzyme transketolase, *Bioorg. Med. Chem. Lett.* 18 (2008) 509–512, <https://doi.org/10.1016/j.bmcl.2007.11.098>.
- A. Iqbal, E.-H. Sahraoui, F.J. Leeper, Gold(I)-catalysed synthesis of a furan analogue of thiamine pyrophosphate, *Beilstein J. Org. Chem.* 10 (2014) 2580–2585, <https://doi.org/10.3762/bjoc.10.270>.
- A.H.Y. Chan, T.C.S. Ho, K. Ageyi-Owusu, F.J. Leeper, Synthesis of pyrrothiamine, a novel thiamine analogue, and evaluation of derivatives as potent and selective inhibitors of pyruvate dehydrogenase, *Org. Biomol. Chem.* 20 (2022) 8855–8858, <https://doi.org/10.1039/D2OB01819E>.
- P.N. Lowe, F.J. Leeper, R.N. Perham, Stereoisomers of Tetrahydrothiamin Pyrophosphate, Potent Inhibitors of the Pyruvate Dehydrogenase Multienzyme Complex from *Escherichia coli*, *Biochemistry* 22 (1983) 150–157, <https://doi.org/10.1021/bi00270a022>.
- J.-B. He, L.-L. Feng, J. Li, R.-J. Tao, Y.-L. Ren, J. Wan, H.-W. He, Design, synthesis and molecular modeling of novel N-acylhydrazine derivatives as pyruvate dehydrogenase complex E1 inhibitors, *Bioorg. Med. Chem.* 22 (2014) 89–94, <https://doi.org/10.1016/j.bmc.2013.11.051>.
- J.-B. He, Y.-L. Ren, Q.-S. Sun, G.-Y. You, L. Zhang, P. Zou, L.-L. Feng, J. Wan, H.-W. He, Design, synthesis and molecular docking of amide and urea derivatives as *Escherichia coli* PDHc-E1 inhibitors, *Bioorg. Med. Chem.* 22 (2014) 3180–3186, <https://doi.org/10.1016/j.bmc.2014.04.003>.
- H. He, H. Xia, Q. Xia, Y. Ren, H. He, Design and optimization of N-acylhydrazine pyrimidine derivatives as E. coli PDHc E1 inhibitors: Structure-activity relationship analysis, biological evaluation and molecular docking study, *Bioorg. Med. Chem.* 25 (2017) 5652–5661, <https://doi.org/10.1016/j.bmc.2017.08.038>.
- Y. Zhou, S. Zhang, H. He, W. Jiang, L. Hou, D. Xie, M. Cai, H. Peng, L. Feng, Design and synthesis of highly selective pyruvate dehydrogenase complex E1 inhibitors as bactericides, *Bioorg. Med. Chem.* 26 (2018) 84–95, <https://doi.org/10.1016/j.bmc.2017.11.021>.
- A.H.Y. Chan, T.C.S. Ho, I. Fathoni, R. Pope, K.J. Saliba, F.J. Leeper, Inhibition of Thiamine Diphosphate (ThDP)-dependent Enzymes by Triazole-based Thiamine Analogues, *ACS Med. Chem. Lett.* (2023), <https://doi.org/10.1021/acsmchemlett.3c00047>.
- L. Feng, J. He, H. He, L. Zhao, L. Deng, L. Zhang, L. Zhang, Y. Ren, J. Wan, H. He, The design, synthesis and biological evaluation of novel thiamin diphosphate analog inhibitors against the pyruvate dehydrogenase multienzyme complex E1 from *Escherichia coli*, *Org. Biomol. Chem.* 12 (2014) 8911–8918, <https://doi.org/10.1039/C4OB01347F>.
- J.-B. He, H.-F. He, L.-L. Zhao, L. Zhang, G.-Y. You, L.-L. Feng, J. Wan, H.-W. He, Synthesis and antifungal activity of 5-iodo-1,4-disubstituted-1,2,3-triazole derivatives as pyruvate dehydrogenase complex E1 inhibitors, *Bioorg. Med. Chem.* 23 (2015) 1395–1401, <https://doi.org/10.1016/j.bmc.2015.02.047>.
- T. Masini, B. Lacy, L. Monjas, D. Hawksley, A.R. de Voogd, B. Illarionov, A. Iqbal, F. J. Leeper, M. Fischer, M. Kontoyianni, A.K.H. Hirsch, Validation of a homology

- model of Mycobacterium tuberculosis DXS: rationalization of observed activities of thiamine derivatives as potent inhibitors of two orthologues of DXS, *Org. Biomol. Chem.* 13 (2015) 11263–11277, <https://doi.org/10.1039/C5OB01666E>.
- [40] L.J.Y.M. Swier, L. Monjas, A. Guskov, A.R. de Voogd, G.B. Erkens, D.J. Slotboom, A.K.H. Hirsch, Structure-based design of potent small-molecule binders to the S-component of the ECF transporter for thiamine, *ChemBiochem* 16 (2015) 819–826, <https://doi.org/10.1002/cbic.201402673>.
- [41] H. He, J. Feng, J. He, Q. Xia, Y. Ren, F. Wang, H. Peng, H. He, L. Feng, Design, synthesis, biological evaluation and molecular docking of amide and sulfamide derivatives as *Escherichia coli* pyruvate dehydrogenase complex E1 inhibitors, *RSC Adv.* 6 (2016) 4310–4320, <https://doi.org/10.1039/C5RA22573F>.
- [42] H. He, W. Wang, Y. Zhou, Q. Xia, Y. Ren, J. Feng, H. Peng, H. He, L. Feng, Rational design, synthesis and biological evaluation of 1,3,4-oxadiazole pyrimidine derivatives as novel pyruvate dehydrogenase complex E1 inhibitors, *Bioorg. Med. Chem.* 24 (2016) 1879–1888, <https://doi.org/10.1016/j.bmc.2016.03.011>.
- [43] L. Monjas, L.J.Y.M. Swier, A.R. de Voogd, R.C. Oudshoorn, A.K.H. Hirsch, D. J. Slotboom, Design and synthesis of thiamine analogues to study their binding to the ECF transporter for thiamine in bacteria, *MedChemComm* 7 (2016) 966–971, <https://doi.org/10.1039/C6MD00022C>.
- [44] Y. Zhou, J. Feng, H. He, L. Hou, W. Jiang, D. Xie, L. Feng, M. Cai, H. Peng, Design, synthesis, and potency of pyruvate dehydrogenase complex E1 inhibitors against cyanobacteria, *Biochemistry* 56 (2017) 6491–6502, <https://doi.org/10.1021/acs.biochem.7b00636>.
- [45] L.J.Y.M. Swier, L. Monjas, F. Reebing, R.C. Oudshoorn, A. Aisyah, T. Primke, M. M. Bakker, E. van Olst, T. Ritschel, I. Faustino, S.J. Marrink, A.K.H. Hirsch, D. J. Slotboom, Insight into the complete substrate-binding pocket of ThiT by chemical and genetic mutations, *MedChemComm* 8 (2017) 1121–1130, <https://doi.org/10.1039/C7MD00079K>.
- [46] L. Monjas, L.J.Y.M. Swier, I. Setyawati, D.J. Slotboom, A.K.H. Hirsch, Dynamic combinatorial chemistry to identify binders of ThiT, an S-component of the energy-coupling factor transporter for thiamine, *ChemMedChem* 12 (2017) 1693–1696, <https://doi.org/10.1002/cmdc.201700440>.
- [47] X.X. Wang, H.Y. Qi, J. Chen, Y.Z. Yang, W. Qiu, W. Wang, P. Zou, B. Li, Y.L. Wang, H.W. He, G.C. Sun, Antibacterial activity and mechanism of ThDP analogs against rice brown stripe pathogen *Acidovorax avenae* subsp. *avenae* RS-1, *J. Plant Pathol.* 101 (2019) 59–69, <https://doi.org/10.1007/s42161-018-0137-4>.
- [48] J. Feng, H. He, Y. Zhou, M. Cai, H. Peng, H. Liu, L. Liu, L. Feng, H. He, Structure optimization and bioactivity evaluation of ThDP analogs targeting cyanobacterial pyruvate dehydrogenase E1, *Bioorg. Med. Chem.* 27 (2019) 115159, <https://doi.org/10.1016/j.bmc.2019.115159>.
- [49] J. He, H. He, M. Cai, F. Zhao, H. He, Insight into the halogen bonding between PA-1 ligand and pyruvate dehydrogenase complex E1 component by crystal structure, DFT calculation, and molecular docking, *J. Molec. Struct.* 1199 (2020), 126991, <https://doi.org/10.1016/j.molstruc.2019.126991>.
- [50] J. Feng, H. He, Y. Zhou, X. Guo, H. Liu, M. Cai, F. Wang, L. Feng, H. He, Design, synthesis and biological evaluation of novel inhibitors against cyanobacterial pyruvate dehydrogenase multienzyme complex E1, *Bioorg. Med. Chem.* 27 (2019) 2413–2420, <https://doi.org/10.1016/j.bmc.2019.01.021>.
- [51] Y. Zhou, J. Feng, L. Feng, D. Xie, H. Peng, M. Cai, H. He, Synthesis and Activity of 1,2,3-Triazole Aminopyrimidines against Cyanobacteria as PDHc-E1 Competitive Inhibitors, *J. Agric. Food Chem.* 67 (2019) 12538–12546, <https://doi.org/10.1021/acs.jafc.9b02878>.
- [52] Y. Zhou, S. Zhang, M. Cai, K. Wang, J. Feng, D. Xie, L. Feng, H. Peng, H. He, Design, Synthesis, and Antifungal Activity of 2,6-Dimethyl-4-aminopyrimidine Hydrazones as PDHc-E1 Inhibitors with a Novel Binding Mode, *J. Agric. Food Chem.* 69 (2021) 5804–5817, <https://doi.org/10.1021/acs.jafc.0c07701>.
- [53] Y. Zhou, M. Cai, H. Zhou, L. Hou, H. Peng, H. He, Discovery of efficient inhibitors against pyruvate dehydrogenase complex component E1 with bactericidal activity using computer aided design, *Pestic. Biochem. Physiol.* 177 (2021), 104894, <https://doi.org/10.1016/j.pestbp.2021.104894>.
- [54] E. Grabowska, M. Czerniecka, U. Czyżewska, A. Zambrzycka, Z. Łotowski, A. Tylicki, Differences in the efficiency of 3-deazathiamine and oxythiamine pyrophosphates as inhibitors of pyruvate dehydrogenase complex and growth of HeLa cells *in vitro*, *J. Enz. Inhib. Med. Chem.* 36 (2021) 122–129, <https://doi.org/10.1080/14756366.2020.1844681>.
- [55] X. Zhang, F. Basuli, S. Abdelwahed, T. Begley, R. Swenson, Radiosynthesis of 5-^[18F]Fluoro-1,2,3-triazoles through Aqueous Iodine-^[18F]Fluorine Exchange Reaction, *Molecules* 26 (2021) 5522, <https://doi.org/10.3390/molecules26185522>.
- [56] D. Zhu, S. Johannsen, T. Masini, C. Simonin, J. Hauptenthal, B. Illarionov, A. Andreas, M. Awale, R.M. Gierse, T. van der Laan, R. van der Vlag, R. Nasti, M. Poizat, E. Buhler, N. Reiling, R. Müller, M. Fischer, J.-L. Reymond, A.K. H. Hirsch, Discovery of novel drug-like antitubercular hits targeting the MEP pathway enzyme DXPS by strategic application of ligand-based virtual screening, *Chem. Sci.* 13 (2022) 10686–10698, <https://doi.org/10.1039/d2sc02371g>.
- [57] A.H.Y. Chan, I. Fathoni, T.C.S. Ho, K.J. Saliba, F.J. Leeper, Thiamine analogues as inhibitors of pyruvate dehydrogenase and discovery of a thiamine analogue with non-thiamine related antiplasmodial activity, *RSC Med. Chem.* 13 (2022) 817–821, <https://doi.org/10.1039/D2MD00085G>.
- [58] A.H.Y. Chan, T.C.S. Ho, D. Parle, F.J. Leeper, Furan-based inhibitors of pyruvate dehydrogenase: SAR study, biochemical evaluation and computational analysis, *Org. Biomol. Chem.* 21 (2023) 1755–1763, <https://doi.org/10.1039/D2OB02272A>.
- [59] Y. Zhou, Y. Qin, H. Zhou, T. Zhang, J. Feng, D. Xie, L. Feng, H. Peng, H. He, M. Cai, Design, synthesis, high algicidal potency, and putative mode of action of new 2-cyclopropyl-4-aminopyrimidine hydrazones, *Pestic. Biochem. Physiol.* 184 (2022) 105098, <https://doi.org/10.1016/j.pestbp.2022.105098>.
- [60] H. Zhao, L.P.S. de Carvalho, C. Nathan, O. Ouerfelli, A protecting group-free synthesis of deazathiamine: A step toward inhibitor design, *Bioorg. Med. Chem. Lett.* 20 (2010) 6472–6474, <https://doi.org/10.1016/j.bmcl.2010.09.053>.
- [61] F. Jordan, H. Li, A. Brown, Remarkable stabilization of zwitterionic intermediates may account for a billion-fold rate acceleration by thiamin diphosphate-dependent decarboxylases, *Biochemistry* 38 (1999) 6369–6373, <https://doi.org/10.1021/bi990373g>.
- [62] D. Barteck, C.L. Freil Meyers, Targeting the unique mechanism of bacterial 1-deoxy-D-xylulose-5-phosphate synthase, *Biochemistry* 57 (2018) 4349–4356, <https://doi.org/10.1021/acs.biochem.8b00548>.
- [63] C.T. Hensley, B. Faubert, Q. Yuan, N. Lev-Cohain, E. Jin, J. Kim, L. Jiang, B. Ko, R. Skelton, L. Loudat, M. Wodzak, C. Klimko, E. McMillan, Y. Butt, M. Ni, D. Oliver, J. Torrealba, C.R. Malloy, K. Kernstine, R.E. Lenkinski, R.J. DeBerardinis, Metabolic heterogeneity in human lung tumors, *Cell* 164 (2016) 681–694, <https://doi.org/10.1016/j.cell.2015.12.034>.
- [64] J. Chen, I. Guccini, D. Di Mitri, D. Brina, A. Revandkar, M. Sarti, E. Pasquini, A. Alajati, S. Pinton, M. Losa, G. Civenni, C.V. Catapano, J. Sgrignani, A. Cavalli, R. D'Antuono, J.M. Asara, A. Morandi, P. Chiarugi, S. Crotti, M. Agostini, M. Montopoli, I. Masgras, A. Rasola, R. Garcia-Escudero, N. Delaleu, A. Rinaldi, F. Bertoni, J. de Bono, A. Carracedo, A. Alimonti, Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer, *Nat. Genet.* 50 (2018) 219–228, <https://doi.org/10.1038/s41588-017-0026-3>.
- [65] R. Hamid, S. Adam, A. Lacour, L.M. Gomez, A.K.H. Hirsch, Structural analysis of 1-deoxy-D-xylulose 5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding, *BioRxiv* (2022), <https://doi.org/10.1101/2022.07.04.498669>.
- [66] E.S. Rudge, A.H.Y. Chan, F.J. Leeper, Prodrugs of pyrophosphates and bisphosphonates: disguising phosphorus oxyanions, *RSC Med. Chem.* 13 (2022) 375–391, <https://doi.org/10.1039/D1MD00297J>.

Strategies on Designing Thiamine Analogues for Inhibition of Thiamine Diphosphate (ThDP)-dependent Enzymes: Systematic Investigation through Scaffold Switching and C2-Functionalisation

Alex H. Y. Chan,^a Terence C. S. Ho,^a Rimsha Irfan,^a Rawia A. A. Hamid,^{bc} Emma S. Rudge,^a Amjid Iqbal,^{ad} Alex Turner,^a Anna K. H. Hirsch,^{bc} and Finian J. Leeper^{a*}

^a Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

^b Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123 Saarbrücken, Germany.

^c Department of Pharmacy, Saarland University, Campus Building E8.1, 66123 Saarbrücken, Germany.

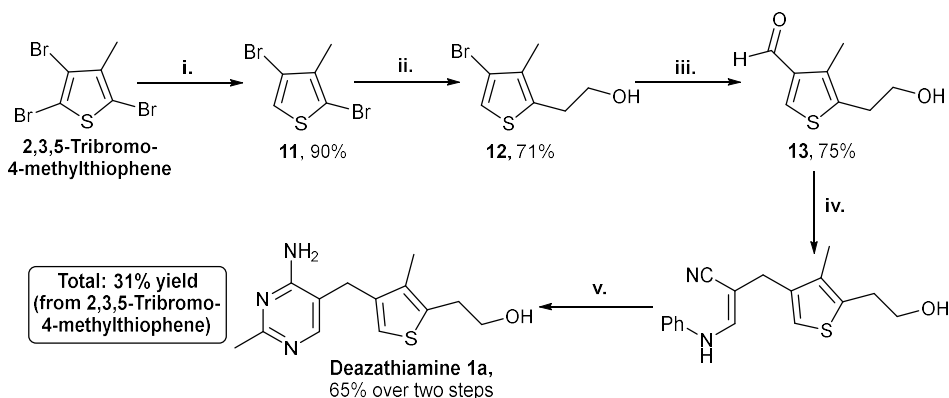
^d Department of Chemistry, Qassim University, Buraydah, Qassim, Saudi Arabia.

* Corresponding author, e-mail: fjl1@cam.ac.uk

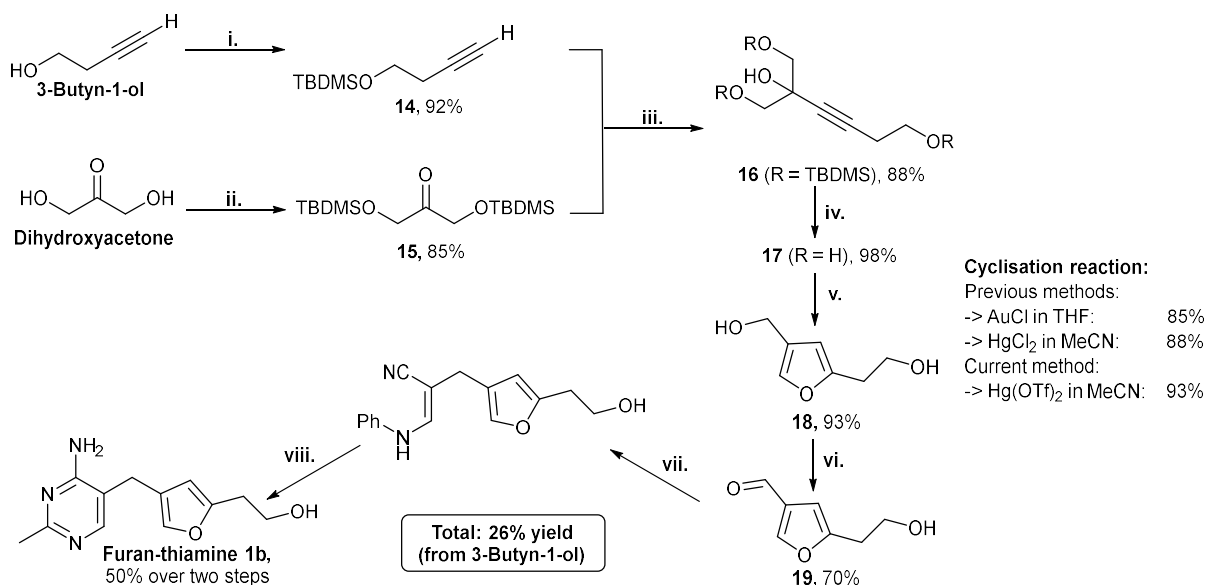
Supplementary Information (SI)

Synthetic Schemes S1–7	S2
Enzyme Inhibition Assays – Methods and Results (no C2-substituent) (Figures S1–4)	S6
Determination of Exit Vector Geometry – Methods and Results (Figure S5)	S13
Computational Docking – Methods and Results (no C2-substituent) (Figure S6)	S14
Enzyme Inhibition Assays –Results (with C2-substituent) (Figures S7–11)	S15
Computational Docking –Results (with C2-substituent) (Figure S12)	S18
Enzyme Inhibition Assays –Results (compound 9) (Figure S13)	S19
Crystallography Data	S20
General Synthesis Methods	S21
Experimental Procedures – Synthesis and NMR spectra	
• Thiamine analogues 1	S21
• ThDP analogues 2	S46
• Esters 3	S56
• ThMP analogues 4	S66
• 5-Acyl deaza-ThMP derivatives 8	S71
• 5-Acetyl deaza-ThDP 9	S96
References	S98

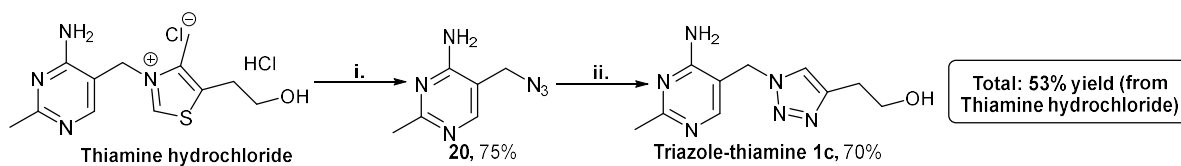
Synthetic Schemes S1-8



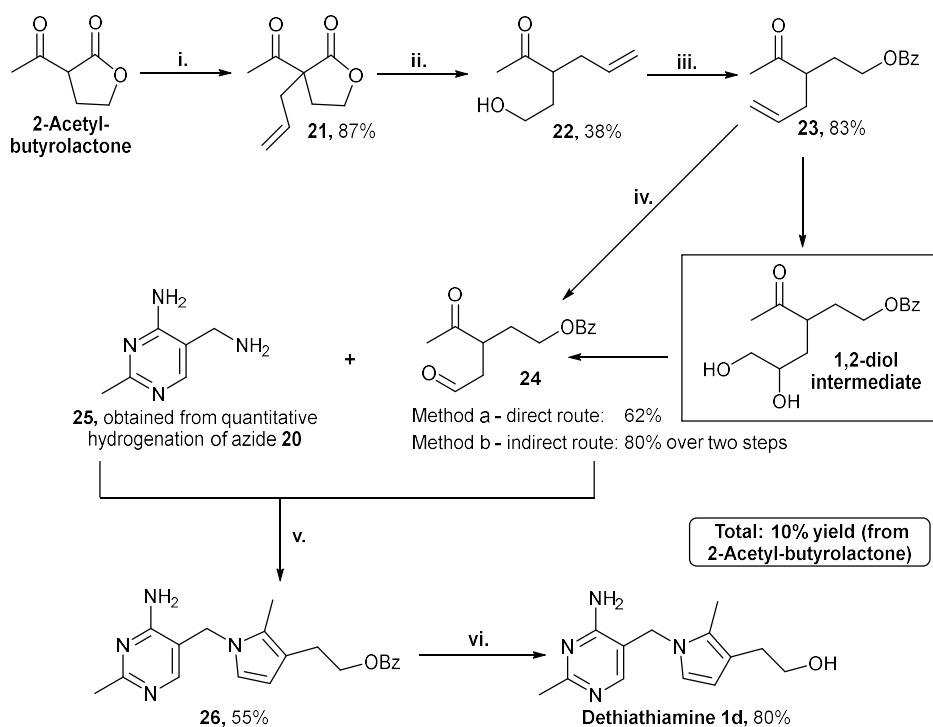
Scheme S1. Synthesis of deazathiamine 1a.¹⁻⁴ *Reagents and conditions:* i) *s*-BuLi, Et₂O, -78 °C, then quenched with MeOH; ii) *n*-BuLi, ethylene oxide, Et₂O·BF₃, Et₂O, RT; iii) *i*-PrMgCl, *n*-BuLi, THF, -40 °C, then DMF, RT; iv) β-anilinopropionitrile, MeONa/MeOH, DMSO, 45 °C; v) acetamidine hydrochloride, EtONa/EtOH, reflux.



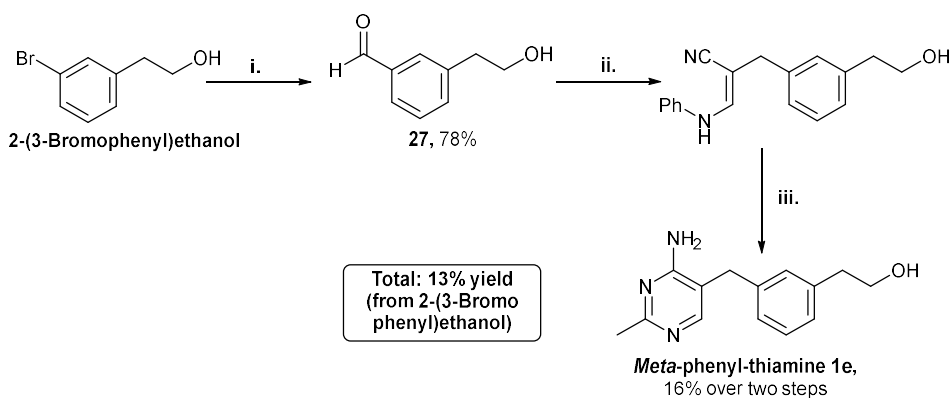
Scheme S2. Synthesis of furan-thiamine 1b.⁵⁻⁸ *Reagents and conditions:* i) TBDMS-Cl, imidazole, DCM, RT; ii) TBDMS-Cl, imidazole, DMF, RT; iii) *n*-BuLi, THF, RT; iv) TBAF, THF, RT; v) Hg(OTf)₂, MeCN, RT; vi) activated MnO₂, CHCl₃, RT; vii) β-anilinopropionitrile, MeONa/MeOH, DMSO, 45 °C; viii) acetamidine hydrochloride, EtONa/EtOH, reflux.



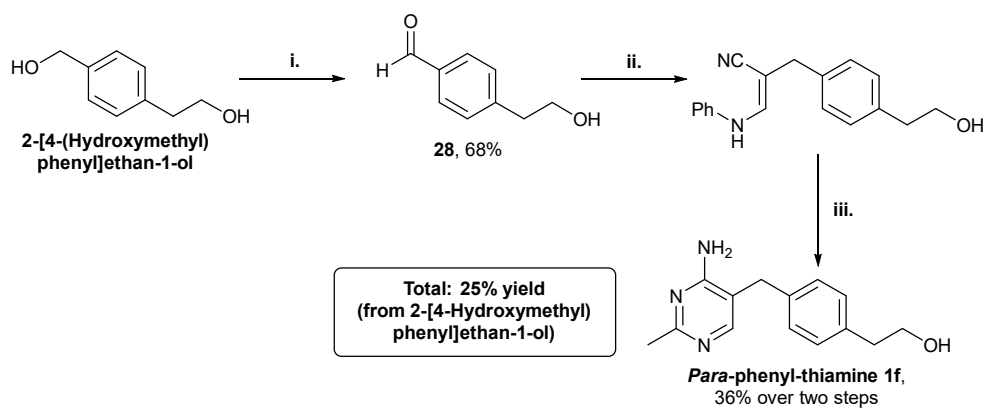
Scheme S3. Synthesis of triazole-thiamine 1c.^{9,10} *Reagents and conditions:* i) NaN₃, Na₂SO₃, H₂O, 65 °C; ii) 3-butyn-1-ol, CuSO₄·5H₂O, sodium ascorbate, *t*-BuOH, H₂O, RT.



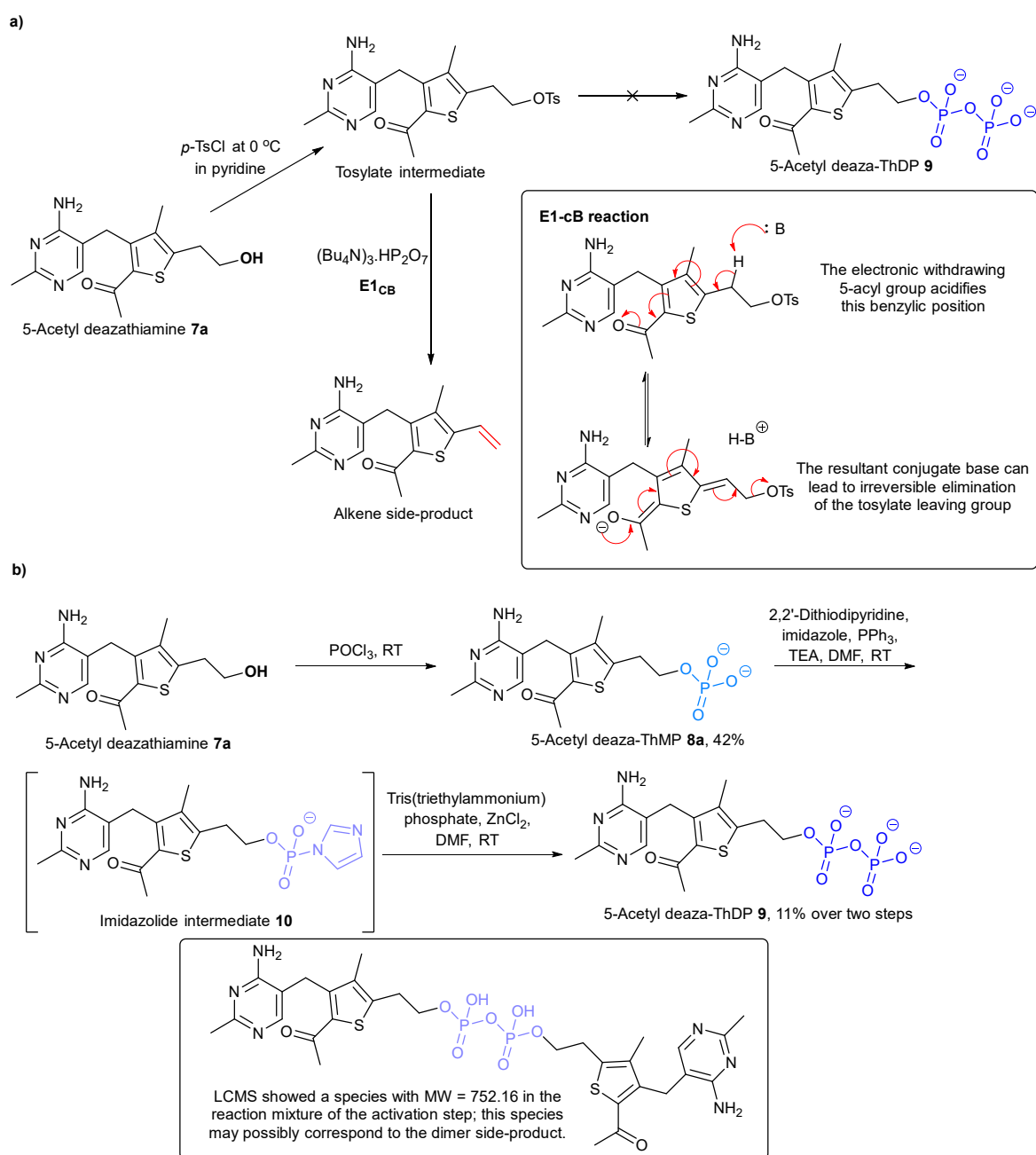
Scheme S4. Synthesis of pyrrothiamine 1d.¹¹ *Reagents and conditions:* i) NaH, THF, allyl bromide, RT; ii) LiOH monohydrate, THF, H₂O, 40 °C; iii) benzoyl chloride, pyridine, 40 °C; iv) Method a: Ozone, zinc dust, AcOH, DCM, RT; Method b: AD-mix-a, methanesulfonamide, *t*-BuOH, H₂O, RT, then, NaIO₄, acetone, H₂O, RT; v) DMF, 45 °C; vi) K₂CO₃, MeOH, 30 °C.



Scheme S5. Synthesis of *meta*-phenyl-thiamine 1e.^{2,4} *Reagents and conditions:* i) *i*-PrMgCl, *n*-BuLi, THF, -40 °C, then DMF, RT; ii) β-anilinopropionitrile, MeONa/MeOH, DMSO, 45 °C; iii) acetamidine hydrochloride, EtONa/EtOH, reflux.



Scheme S6. Synthesis of *para*-phenyl-thiamine 1f. Reagents and conditions: i) LiAlH_4 , THF, 50 °C; ii) activated MnO_2 , CHCl_3 , RT; iii) β -anilinopropionitrile, MeONa/MeOH , DMSO, 45 °C; iv) acetamidine hydrochloride, EtONa/EtOH , reflux.



Scheme S7. a) Attempted synthesis of 5-acetyl deaza-ThDP 9 from a tosylate intermediate. Preparing 5-acetyl deaza-ThDP derivatives by this scheme failed as the tosylation conditions favours the E1-cB elimination reaction of the tosylate species. **b) Preparation of 9 from an imidazolide intermediate 10.** The monophosphate precursor **8a** was activated and then displaced by a nucleophilic phosphorus species. The yield was poor, and several issues associated with the activation step were then identified: 1) the conversion rate of the monophosphate to the activated imidazolide was low; 2) the reactivity of the imidazolide species resulted in hydrolytic degradation (back to the starting material) in the work-up procedures; 3) some unreacted **8a** coupled to the imidazolide intermediate to form a dimer.

Enzyme Inhibition Assays – Methods and Results (Figures S1–4, S7–11 and S13)

Evaluation of the inhibitory activity of compounds against Porcine PDH E1 *in vitro*

Porcine PDH E1 was purchased from Sigma. Porcine PDH E1 activity was determined by monitoring 2,6-dichlorophenolindophenol (DCPIP) reduction at 600 nm using a microplate reader (CLARIOstar) and conducted as our previously described methods with some modifications.⁸ The reaction buffer (50 mM KH₂PO₄ and 1 mM MgCl₂, pH 7) contained thiamine pyrophosphate (ThDP) at 10 μM, 0.25 mM 2,6-dichlorophenolindophenol (DCPIP), and 2 mg/ml porcine PDH E1. The reaction mixture was preincubated with inhibitors at 37 °C for 30 min, then reaction was initiated by adding pyruvate to a final concentration of 50 mM. The percentage inhibition of compounds against porcine PDH E1 was assayed at [ThDP]:[inhibitor] of 1:1 (10:10 μM) and 20:1 (200:10 μM); percentage inhibition of **4a-c**, **e**, and **f** dropped from 90-63% (10:10 μM) to 41-7% (200:10 μM). To determine the half-maximal inhibitory concentration (IC₅₀), ThDP concentration was set at 10 μM (with the exception of testing **2a-f** where ThDP was set at 60 μM) with varying inhibitor concentrations. Specific activity was calculated using the molar extinction coefficient of DCPIP, 21 mM⁻¹ cm⁻¹.^{12,13} K_M of ThDP was reported to be 0.05 μM.¹⁴ The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

Evaluation of the inhibitory activity of compounds against *S. cerevisiae* PDC *in vitro*

S. cerevisiae PDC was purchased from Sigma. *S. cerevisiae* PDC activity was determined by monitoring 2,6-dichlorophenolindophenol (DCPIP) reduction at 600 nm using a microplate reader (CLARIOstar) and conducted as described above with some modifications.⁸ The reaction buffer (50 mM KH₂PO₄ and 1 mM MgCl₂, pH 7) contained ThDP at 150 μM, 0.27 mM DCPIP, and 0.15 mg/ml *S. cerevisiae* PDC. The reaction mixture was preincubated at 37 °C with inhibitors for 60 min, then reaction was initiated by adding pyruvate to a final concentration of 70 mM. The percentage inhibition of compounds against *S. cerevisiae* PDC was assayed at [ThDP]:[inhibitor] of 1:1 (150:150 μM) and 20:1 (3000:150 μM); percentage inhibition of **4a-c**, **e**, and **f** dropped from 32-11% (150:150 μM) to <4% (3000:150 μM). To determine the half-maximal inhibitory concentration (IC₅₀), ThDP concentration was set at 150 μM with varying inhibitor concentrations. Specific activity was calculated using the molar extinction coefficient of DCPIP, 21 mM⁻¹ cm⁻¹.^{12,13} K_M of ThDP was found to be 15 μM with the assay conditions above in the absence of inhibitor compounds and varying ThDP concentration. The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

Evaluation of the inhibitory activity of compounds against *Z. mobilis* PDC *in vitro*

Z. mobilis PDC was expressed and purified following a published method.^{7,8} *Z. mobilis* PDC activity was determined by monitoring reduced nicotinamide adenine dinucleotide (NADH) consumption at 340 nm using a microplate reader (CLARIOstar) and conducted as described by with some modifications.^{7,8} The reaction buffer (50 mM MES-KOH and 5 mM MgCl₂, pH 6.5) contained ThDP at 10 μM, 150 μM NADH, 10 U/ml alcohol dehydrogenase (ADH) and 0.5 μM of active sites of *Z. mobilis* PDC. To determine the half-maximal inhibitory concentration (IC₅₀), ThDP concentration was set at 10 μM with varying inhibitor concentrations. The reaction mixture was preincubated at 37 °C with inhibitors for 60 min, then reaction was initiated by adding pyruvate to a final concentration of 10 mM. K_M of ThDP was reported to be 0.35 μM.¹⁵ The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

Evaluation of the inhibitory activity of compounds against *E. coli* OGDH E1 *in vitro*

E. coli OGDH E1 was from material donated by R. Frank. Activity was determined by monitoring 2,6-dichlorophenolindophenol (DCPIP) reduction at 600 nm using a microplate reader (CLARIOstar) and conducted as described with some modifications.^{2,8} The reaction buffer (50 mM KH₂PO₄ and 2 mM MgCl₂, pH 7) contained ThDP at 30 μM, 0.5 mM DCPIP, and 6.7 mg/ml *E. coli* OGDH E1. The reaction mixture was preincubated at 37 °C with inhibitors for 60 min, then reaction was initiated by adding α-ketoglutarate to a final concentration of 10 mM. The percentage inhibition of compounds against *E. coli* OGDH E1 was assayed at [ThDP]:[inhibitor] of 1:1 (30:30 μM) and 20:1 (600:30 μM); percentage inhibition of **4a-c**, **e**, and **f** dropped from 53-19% (30:30 μM) to <5% (600:30 μM). To determine the half-maximal inhibitory concentration (IC₅₀), ThDP concentration was set at 30 μM (with the exception of testing **2a-f** where ThDP was set at 60 μM) with varying inhibitor concentrations. Specific activity was calculated using the molar extinction coefficient of DCPIP, 21 mM⁻¹ cm⁻¹.^{12,13} K_M of ThDP was reported to be 3 μM.² The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

Evaluation of the inhibitory activity of compounds against *A. viridans* Pyruvate Oxidase *in vitro*

A. viridans pyruvate oxidase and horseradish peroxidase were purchased from Sigma. *A. viridans* pyruvate oxidase activity was determined by monitoring appearance of quinoneimine dye at 550 nm using a microplate reader (CLARIOstar) and conducted as described with some modifications.^{8,16} The reaction buffer (50 mM KH₂PO₄ and 10 mM MgCl₂, pH 5.9) contained ThDP at 60 μM, 10 μM flavin adenine dinucleotide (FAD), 0.15% 4-aminoantipyrine, 0.3% N-ethyl-N-(2-hydroxy-3-sulfopropyl)-m-toluidine (EHSPT), 50 μg/mL horseradish peroxidase and 0.35 U/mL *A. viridans* pyruvate oxidase. The reaction mixture was preincubated at 37 °C with inhibitors for 30 min, then reaction was initiated by adding pyruvate to a final concentration of 50 mM. The percentage inhibition of compounds against *A. viridans* pyruvate oxidase was assayed at [ThDP]:[inhibitor] of 1:1 (60:60 μM) and 20:1 (1200:60 μM); percentage inhibition of **4a-c**, **e**, and **f** dropped from 74-22% (60:60 μM) to <10% (1200:60 μM). To determine the half-maximal inhibitory concentration (IC₅₀), ThDP concentration was set at 60 μM with varying inhibitor concentration. 1 unit of pyruvate oxidase activity is defined as the production of 1 μmol of hydrogen peroxide per minute. K_M of ThDP was reported to be 5 μM.¹⁰ The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

Evaluation of the inhibitory activity of compounds against DXP synthase (DXPS) *in vitro*

DXPS and DXP reductoisomerase were obtained from *K. pneumoniae* and *E. coli* respectively. DXPS activity was determined by monitoring consumption of NADPH at 365 nm using a microplate reader (CLARIOstar) and conducted as described with some modifications.^{17,18} The reaction mixture contained 40 mM Tris-HCl (pH 7.5), 0.3 mM NADPH, 1 mM MgCl₂, 0.5 μM ThDP, 6 mM DL-glyceraldehyde 3-phosphate, 3 mM pyruvate, 1 mM 2-mercaptoethanol, 100 mM NaCl, 27 μg (=1 U) of *E. coli* DXP reductoisomerase and 3 μg of DXPS from *K. pneumoniae* in a total volume of 600 μL. To determine the half-maximal inhibitory concentration (IC₅₀), reaction mixtures were preincubated with varying inhibitor concentrations for 5 min at 37 °C and started by addition of pyruvate. K_M of ThDP was reported to be 0.05 μM.¹⁹ The enzyme IC₅₀ values were calculated from non-linear regression curve fitting using GraphPad Prism.

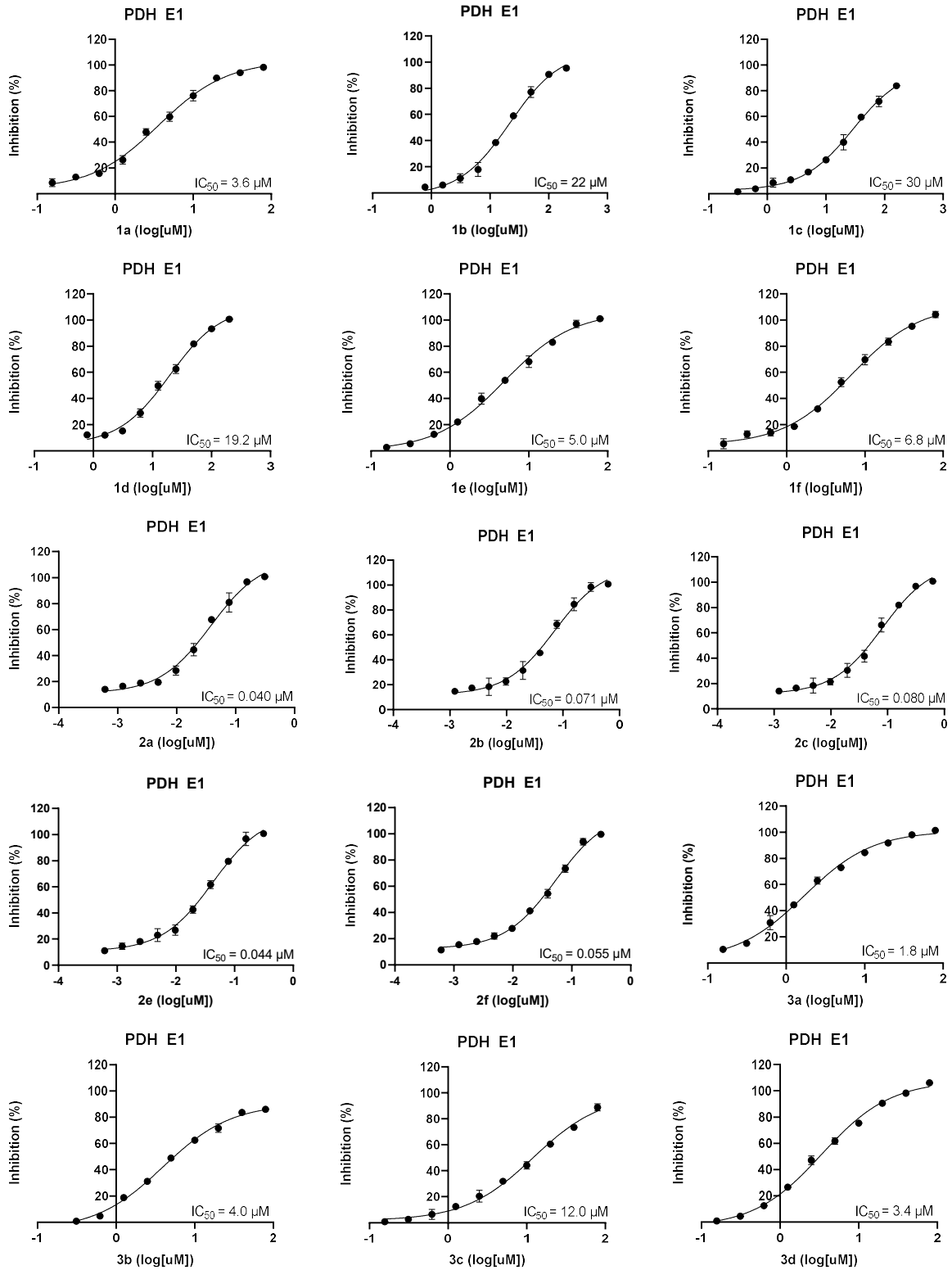


Figure S1 continues on next page

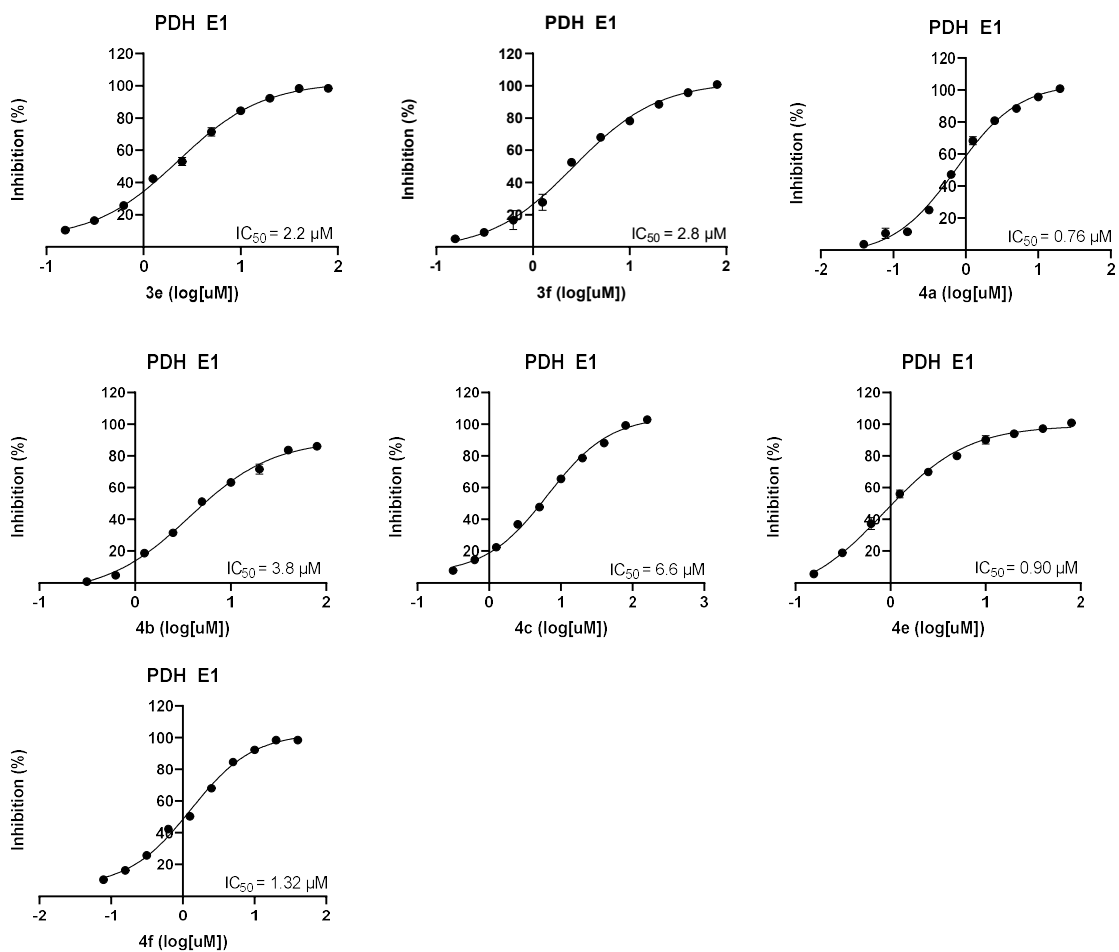


Figure S1. IC_{50} values for compounds (labelled in x-axis) on PDH E1. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

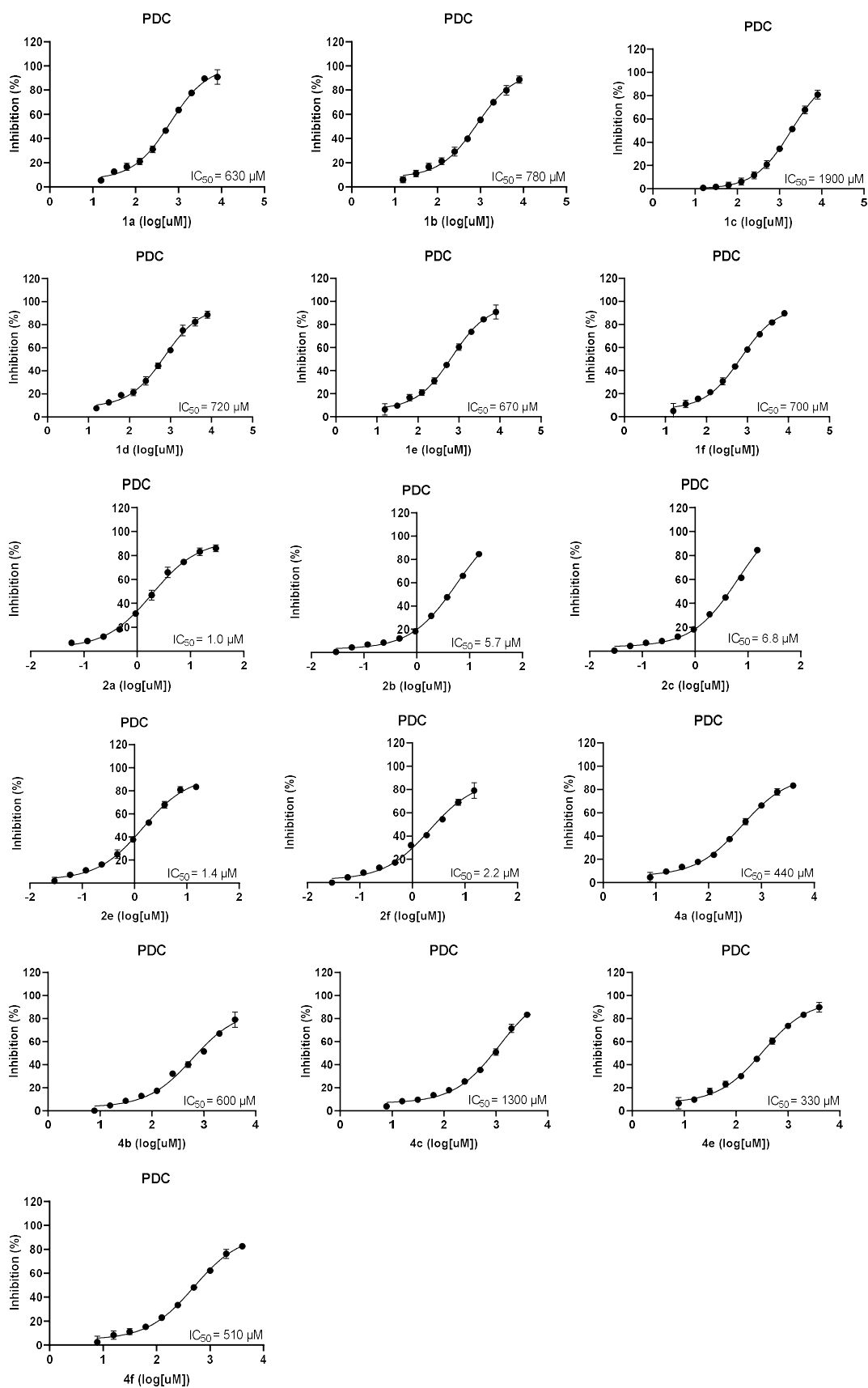


Figure S2. IC_{50} values for compounds (labelled in x-axis) on yeast PDC. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

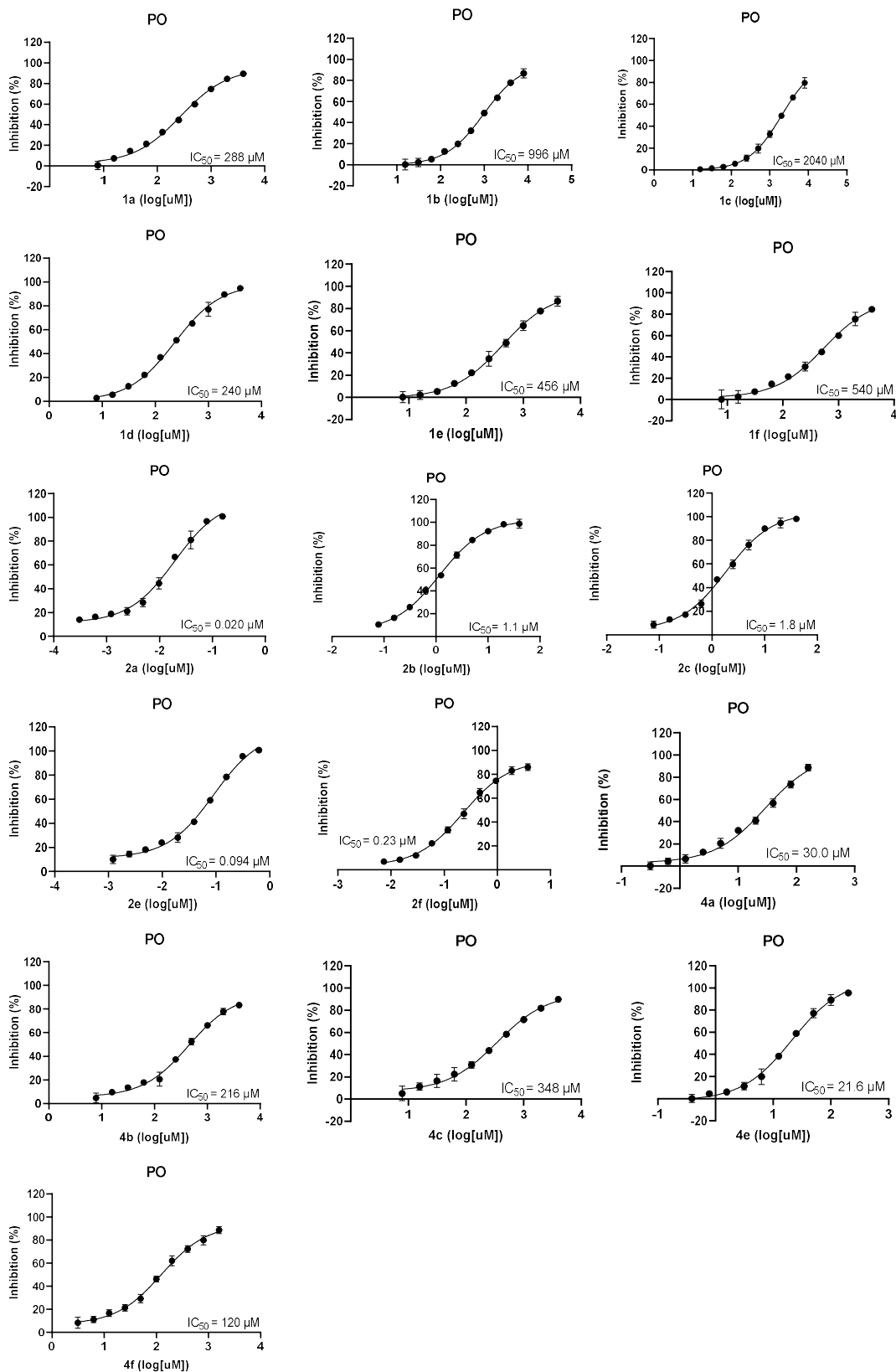


Figure S3. IC₅₀ values for compounds (labelled in x-axis) on *A. viridans* PO. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

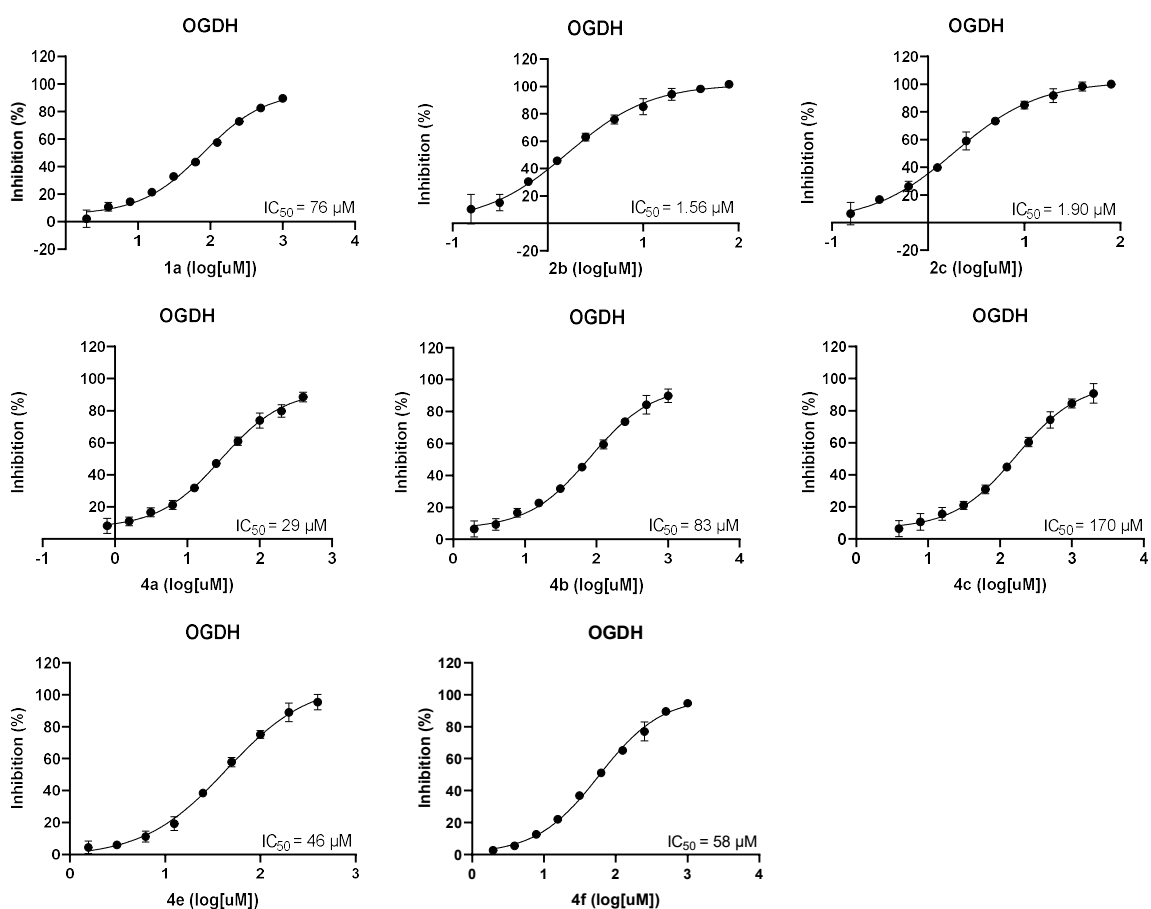


Figure S4. IC_{50} values for compounds (labelled in x-axis) on *E. coli* OGDH. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

Determination of Substituent Angles – Methods and Results (Figure S5)

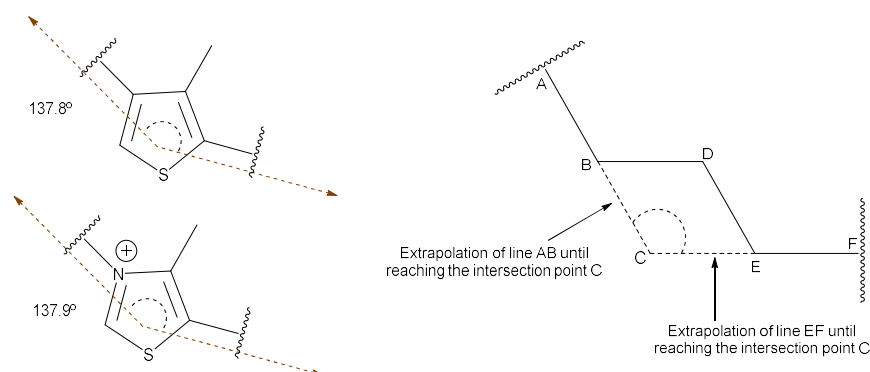


Figure S5. General approach to determining $\angle BCE$ which corresponds to the exit vector geometry of the disubstituted five-membered aromatic rings. Firstly, CDCC Mogul 1.8.4 was used to display the structure of deazathiamine and thiamine and measure $\angle ABD$, $\angle DEF$ and $\angle BDE$. Then $\angle BCE = \angle ABD + \angle DEF - \angle BDE$.

For the thiophene ring (deazathiamine):

The Refcode of deazathiamine in CDCC Mogul 1.8.4 is XEMDAU.¹

Refcode: XEMDAU Data Library: CSD 5.41	
Identifier	XEMDAU
Literature Reference	D.Hawksley, D.A.Griffin, F.J.Leeper, <i>J.Chem.Soc.,Perkin Trans.1</i> (2001), 144, doi: 10.1039/b006962k
Formula	C ₁₃ H ₁₇ N ₃ O S,0.5(H ₂ O)
Compound Name	2-Hydroxyethyl-3-methyl-4-((4'-amino-2-methyl-5-pyrimidyl)methyl)-thiophene hemihydrate
Synonym	Deazathiamine hemihydrate; PDB Chemical Component code: OYN
Space Group	C 2/c (15)
Cell Lengths	a 47.033(2) b 11.2180(7) c 10.5570(5)
Cell Angles	α 90 β 92.501(4) γ 90
Cell Volume	5564.74
Z, Z'	Z: 16 Z': 2
R-Factor (%)	8.51
Disorder	The hydroxyethyl group is disordered over two sites with occupancies 0.67:0.33.
Polymorph	

$\angle ABD$ (at C3) = 122.1°, $\angle DEF$ (at C5) = 128.6° and $\angle BDE = 112.9^\circ$.

So $\angle BCE = 122.1^\circ + 128.6^\circ - 112.9^\circ = 137.8^\circ$

For the thiazolium ring (thiamine):

The Refcode of thiamine in CDCC Mogul 1.8.4 is GEYXOX.²⁰

Refcode: GEYXOX Data Library: CSD 5.41	
Identifier	GEYXOX
Literature Reference	D.A.Clemente, A.Marzotto, G.Valle, <i>J.Crystallogr.Spectrosc.Res.</i> (1988), 18 , 147, doi: 10.1007/BF01181906
Formula	C ₁₂ H ₁₇ N ₄ O S ⁺ ,Cl ⁻
Compound Name	Thiamine chloride
Synonym	
Space Group	P 2 ₁ /a (14)
Cell Lengths	a 18.929(4) b 11.663(2) c 6.376(2)
Cell Angles	α 90 β 96.72(8) γ 90
Cell Volume	1397.95
Z, Z'	Z: 4 Z': 1
R-Factor (%)	4.21
Disorder	Hydrogen atoms of C17 are disordered; only three retained.
Polymorph	

$\angle ABD$ (at N3) = 121.4°, $\angle DEF$ (at C5) = 128° and $\angle BDE = 111.5^\circ$.

So $\angle BCE = 121.4^\circ + 128^\circ - 111.5^\circ = 137.9^\circ$

Computational Docking – Methods and Results

Docking of compounds were executed using CCDC GOLD docking program with PDB: 6CFO, 1VPD and 1V5F for human PDH E1, *Saccharomyces cerevisiae* PDC (ScPDC) and *Aerococcus viridans* PO (AvPO) respectively. The ThDP or equivalent ligand were selected as the binding site. Our molecules were generated using Mercury. GA runs were set at 30 and was user defined with population size of 200 and 200000 number of operations. No early termination was permitted. Scaffold constraint to the original ligand was implemented on our compounds to mimic their binding positions. CHEMPLP and GoldScore were the docking scoring and rescoring respectively. For all other GOLD-specific docking options the default settings were used.²¹

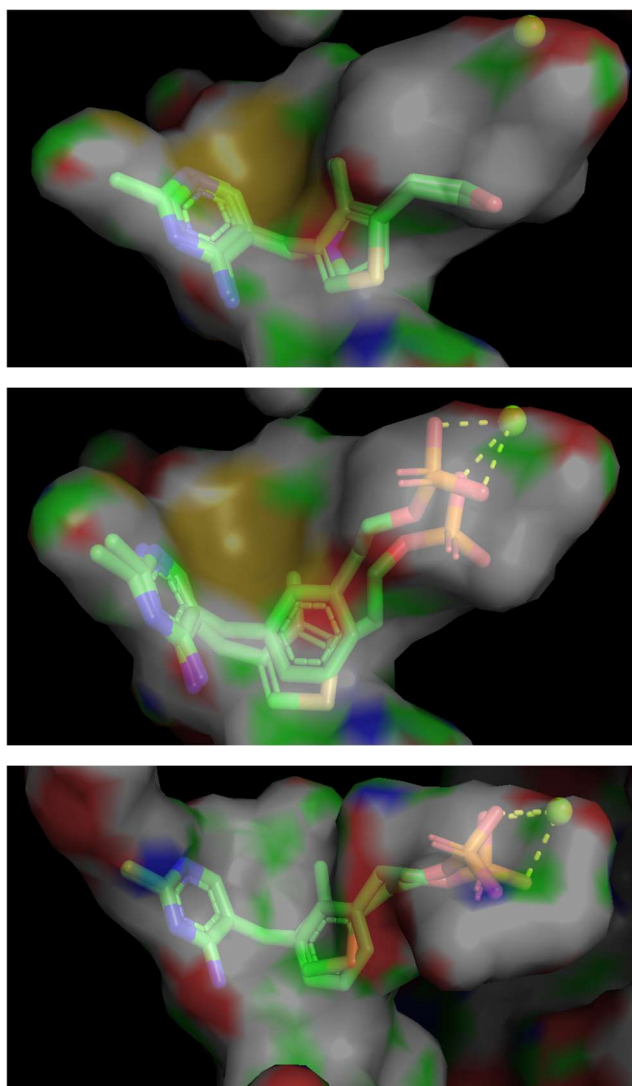


Figure S6. Molecular docking of 1a and 1d in AvPO (PDB: 1V5F) (top), 4a and 4e in AvPO (PDB: 1V5F) (middle) and 4a and 4e in ScPDC (PDB: 1PVD) (bottom); Mg²⁺ shown as a yellow-green sphere; ThDP-binding pocket shown as surface. The experimentally determined affinity of 1a and 1d on PO was too close to be explained by computational docking as the suggested binding poses were almost identical. In the cases of PO and PDC, molecular modelling suggested that the phosphate moiety of 4e (bidentate) was placed closer to the Mg²⁺ than that of 4a (monodentate) possibly due to the different substituent angles of 4a and 4e.

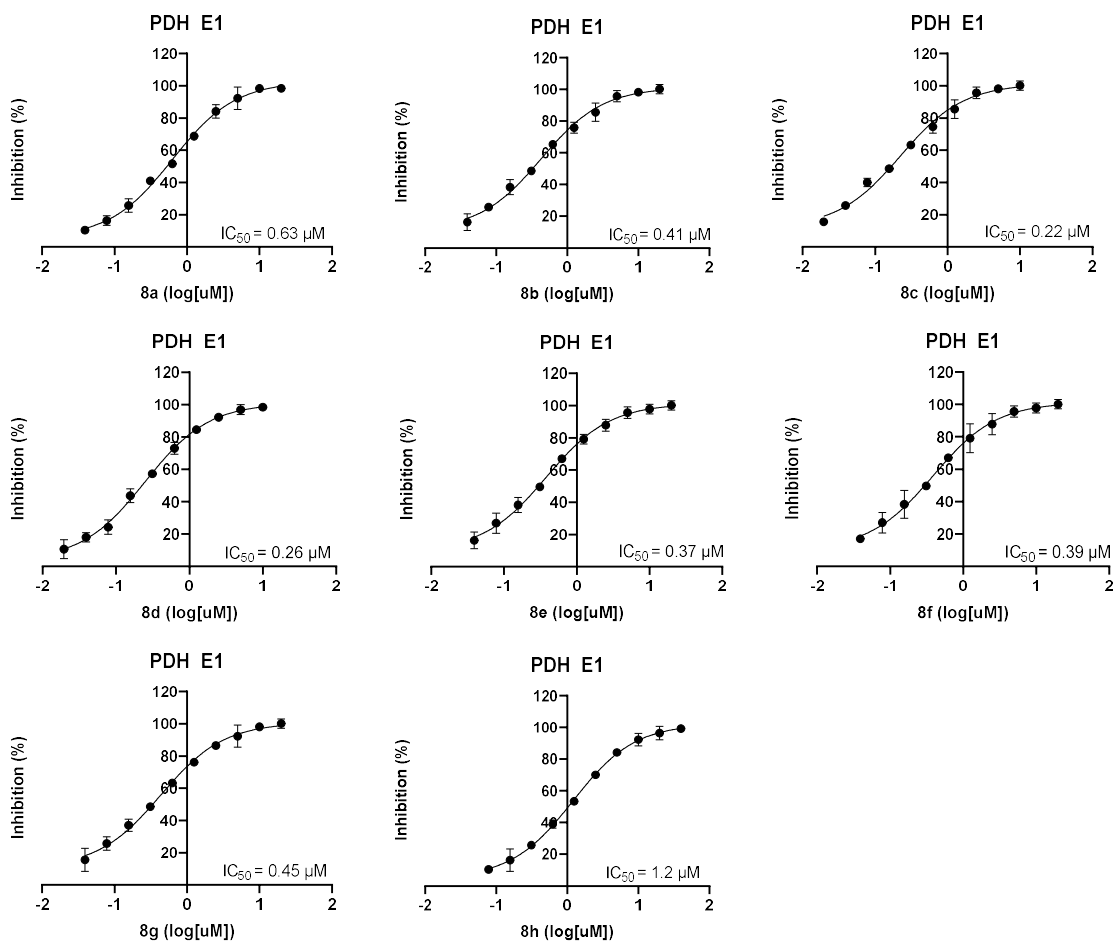


Figure S7. IC₅₀ values for compounds (labelled in x-axis) on PDH E1. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

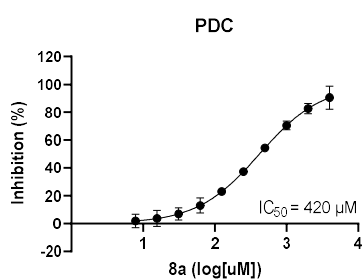


Figure S8. IC₅₀ values for compounds (labelled in x-axis) on yeast PDC. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

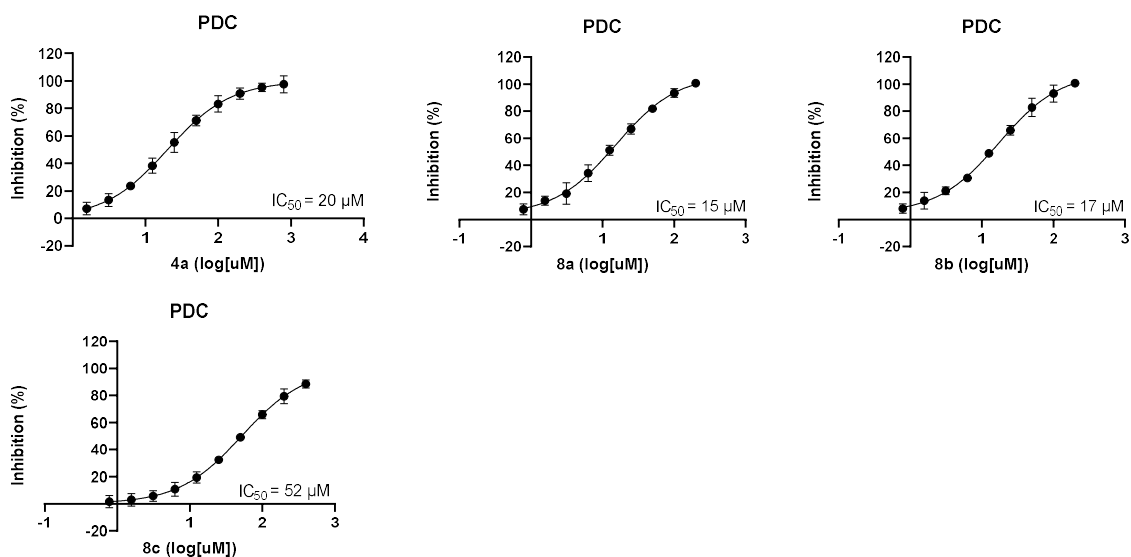


Figure S9. IC₅₀ values for compounds (labelled in x-axis) on bacterial PDC. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

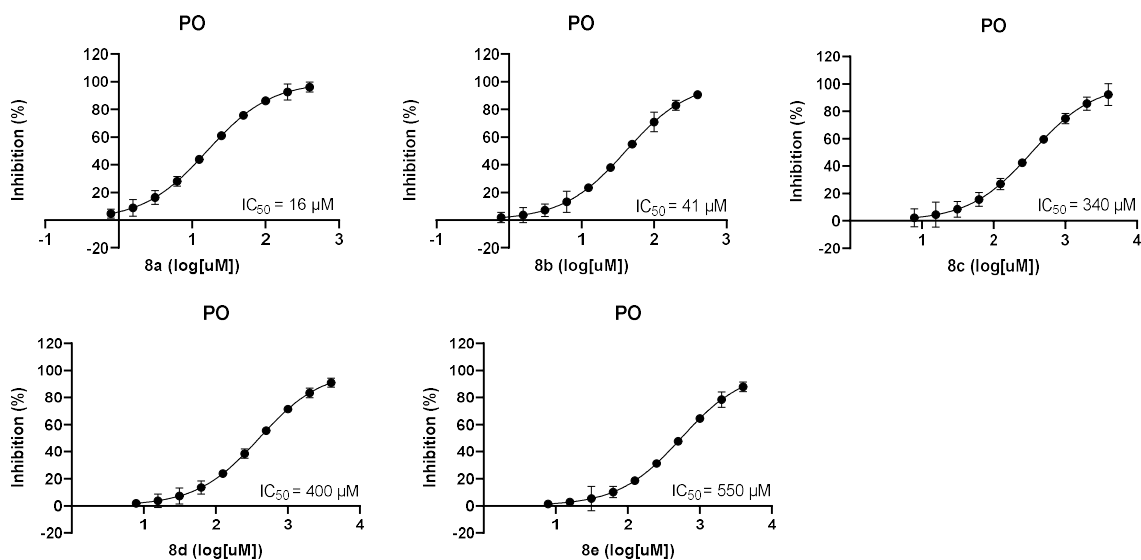


Figure S10. IC₅₀ values for compounds (labelled in x-axis) on *A. viridans* PO. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

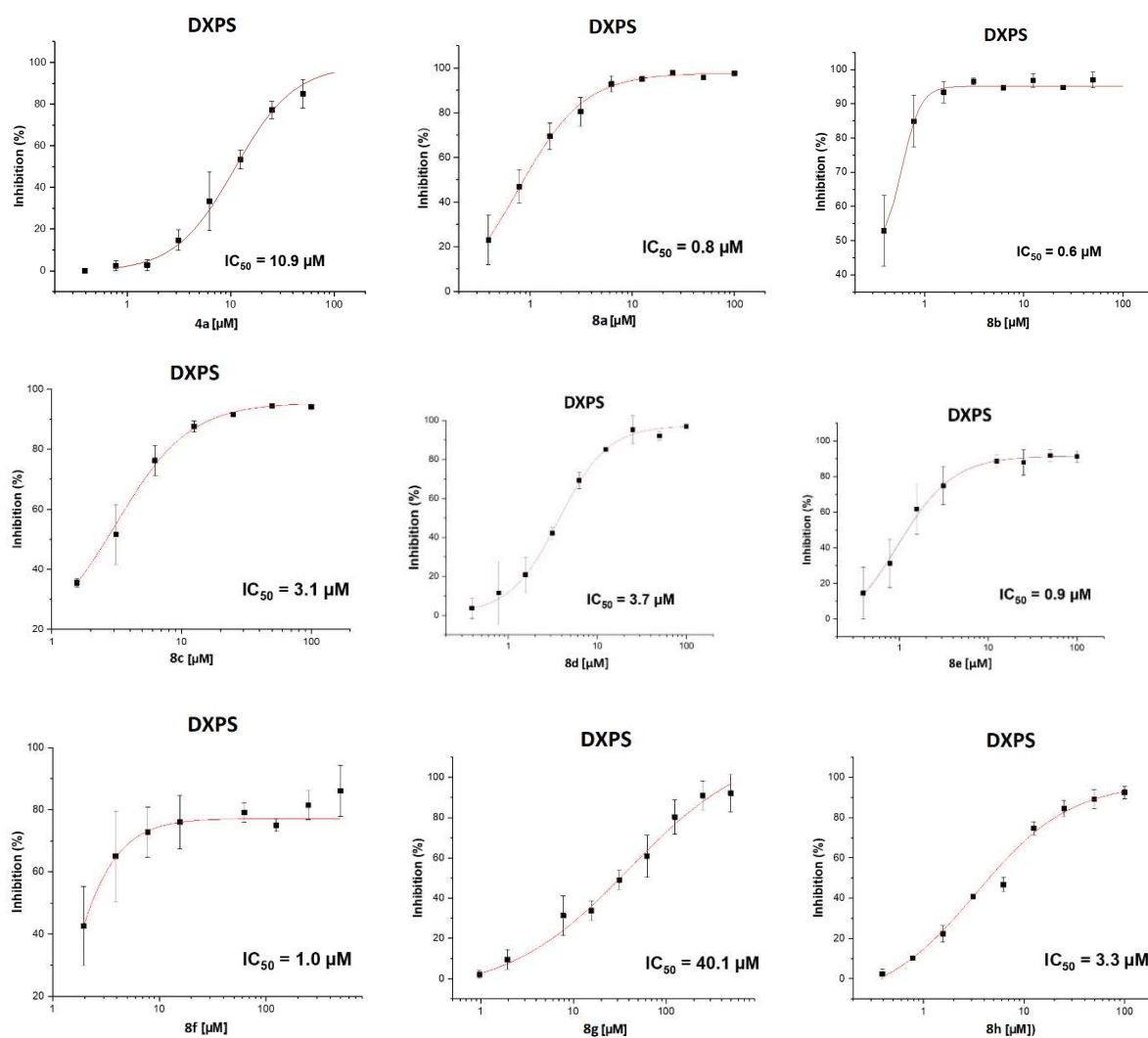


Figure S11. IC_{50} values for compounds (labelled in x-axis) on DXPS. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

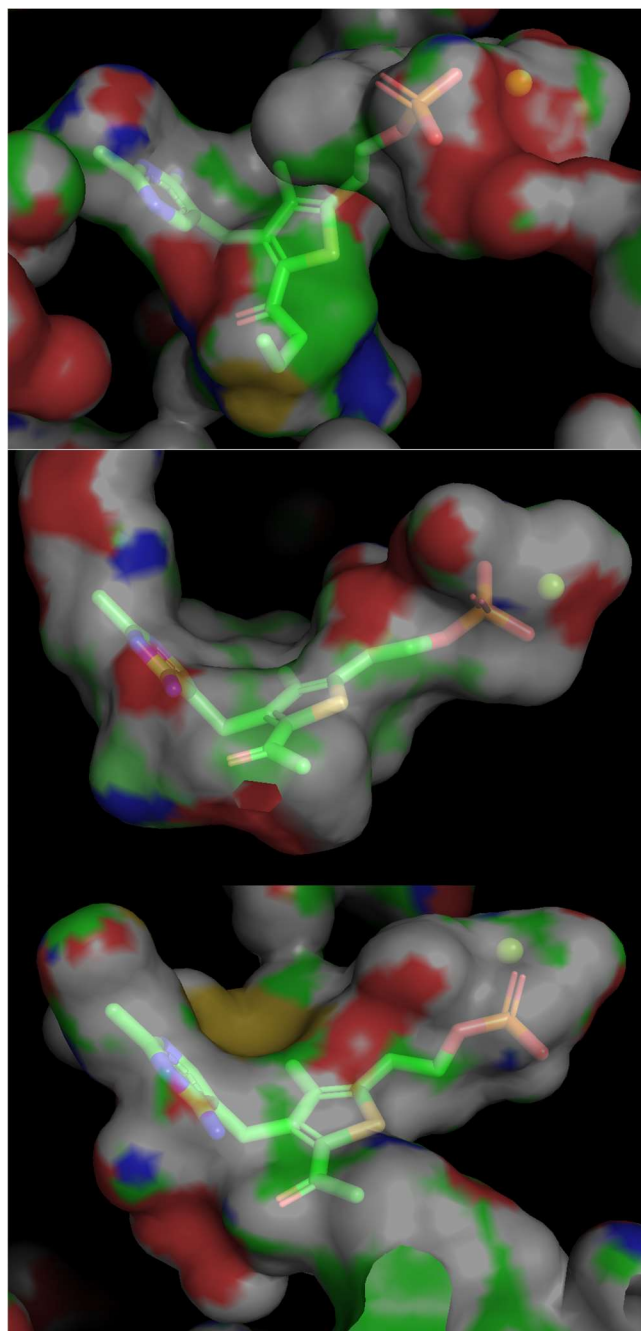


Figure S12. Molecular docking of 8c in human PDH E1 (PDB: 6CFO) (top), 8a in ScPDC (PDB: 1PVD) (middle) and 8a in AvPO (PDB: 1V5F) (bottom); Mg²⁺ shown as a yellow-green sphere; ThDP-binding pocket shown as surface. In all cases, the C2-substituent was shown to extend into the substrate pocket.

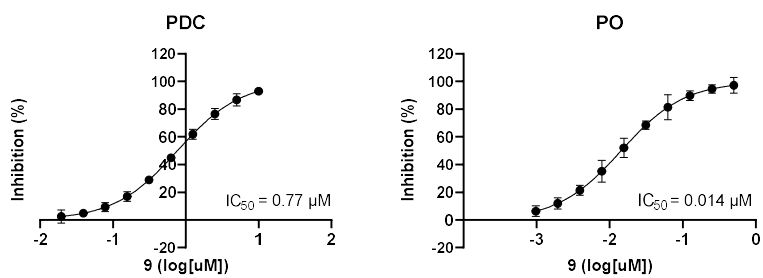


Figure S13. IC₅₀ values for compound 9 on yeast PDC and *A. viridans* PO. Measurements were made in three technical replicates under assay conditions as described. Where the error bars are not visible, they are smaller than the symbols. Best-fit nonlinear regression curves are shown.

Crystallography Data

Data collection and refinement statistics of *KpDXPS* co-crystal with monophosphate **8b**

PDB code	8BZX
Resolution range	46.79 - 2.05 (2.123 - 2.05)
Space group	C 2 2 21
Unit cell	119.249 150.969 142.145 90 90 90
Unique reflections	80381 (7953)
Completeness (%)	99.84 (99.40)
Mean I/sigma(I)	1.6
Wilson B-factor	34.97
R-merge	0.069
Reflections used in refinement	80274 (7908)
Reflections used for R-free	3953 (391)
R-work	0.2170 (0.4197)
R-free	0.2624 (0.4367)
Number of non-hydrogen atoms	8770
macromolecules	8485
ligands	95
solvent	230
Protein residues	1113
RMS(bonds)	0.007
RMS(angles)	0.84
Ramachandran favored (%)	95.73
Ramachandran allowed (%)	3.63
Ramachandran outliers (%)	0.64
Rotamer outliers (%)	2.48
Clashscore	9.70
Average B-factor	49.17
macromolecules	49.34
ligands	42.99
solvent	44.27
Number of TLS groups	1

*Statistics for the highest-resolution shell are shown in parentheses.

General Synthesis Methods

Oxygen- and moisture-sensitive reactions were carried out in flame-dried glassware under a nitrogen atmosphere. Unless otherwise stated, all chemicals and reagents were purchased from commercial suppliers and used without further purification. Reaction progress was monitored by analytical thin-layer chromatography (TLC). TLC was conducted using Merck glass plates with silica Kieselgel 60 F254 of thickness 0.25 mm and visualised under 254 nm UV lamp or potassium permanganate staining solution (with gentle heating). Flash column chromatography was carried out in the indicated solvent system using prepacked silica gel cartridges for use on the Biotage Purification System. All solvents were removed under reduced pressure using a Büchi rotary evaporator with dry ice traps. Reverse-phase HPLC was used to purify mono-phosphorylated compounds: 0-100% MeOH in water (with 0.1% formic acid) and di-phosphorylated compounds: 0-100% MeOH in water (with 10 mM ammonium bicarbonate).

All yields refer to chromatographically and spectroscopically pure compounds unless otherwise stated. Known compounds were characterised by ^1H NMR spectroscopy and ^{13}C NMR spectroscopy. New synthetic intermediates were characterised by, at minimum, ^1H NMR spectroscopy, ^{13}C NMR spectroscopy and ESI-MS unless otherwise stated. Compounds subjected to biological assays were characterised by, at minimum, ^1H NMR spectroscopy, ^{13}C NMR spectroscopy and HRMS.

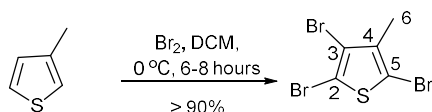
^1H NMR spectra were recorded at 400 MHz or 700 MHz in CDCl_3 , D_2O , CD_3OD or CD_3SOCD_3 solution on a Bruker 400 MHz or 700 MHz spectrometer and chemical shifts were recorded in parts per million (ppm). ^{13}C NMR spectra were recorded on either a Bruker 400 MHz or 700 MHz spectrometer. ^{31}P NMR spectra were recorded on either a Bruker 400 MHz or 700 MHz spectrometer. Resonances are described using the following abbreviations: s (singlet), d (doublet), t (triplet), q (quartet), qn (quintet), sext (sextet), m (multiplet), br (broad), etc. Coupling constants (J) are given in Hz and are rounded to the nearest 0.1 Hz. All NMR data were collected at 25 °C. Mass spectra used electrospray ionisation (ESI). Melting points of compounds were measured using a Reichert machine and are uncorrected.

Experimental Procedures – Synthesis and NMR spectra

Preparation of deazathiamine 1a:

2,3,5-Tribromo-4-methylthiophene

Despite being commercially available, it is cheaper to prepare from 3-methylthiophene:

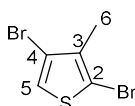


To a stirred solution of 3-methylthiophene (6.48 g, 66 mmol) in DCM (77 mL, 0.85 M) at $0\text{ }^\circ\text{C}$ under dark conditions was added Br_2 (11.8 mL, 231 mmol) dropwise. The resultant mixture was stirred at $0\text{ }^\circ\text{C}$ and the flask was covered with Aluminium foil for 6-8 h. The reaction mixture was quenched with sat. aq. K_2CO_3 (250 mL) slowly at $0\text{ }^\circ\text{C}$ and then extracted with DCM (3 x 300 mL). The combined organic phases were dried over MgSO_4 , filtered, and evaporated under reduced pressure to yield 2,3,5-tribromo-4-methylthiophene as a thick yellow oil/semisolid (21 g, 95%), which was used in the next step without further purification.

^1H NMR (400 MHz, CDCl_3) δ 2.26 (s, 3H, H-5).

^{13}C NMR (100 MHz, CDCl_3) δ 137.2 (C-4), 116.2 (C-3), 109.1 (C-5), 107.9 (C-2), 16.8 (C-5).

2,4-Dibromo-3-methylthiophene **11**



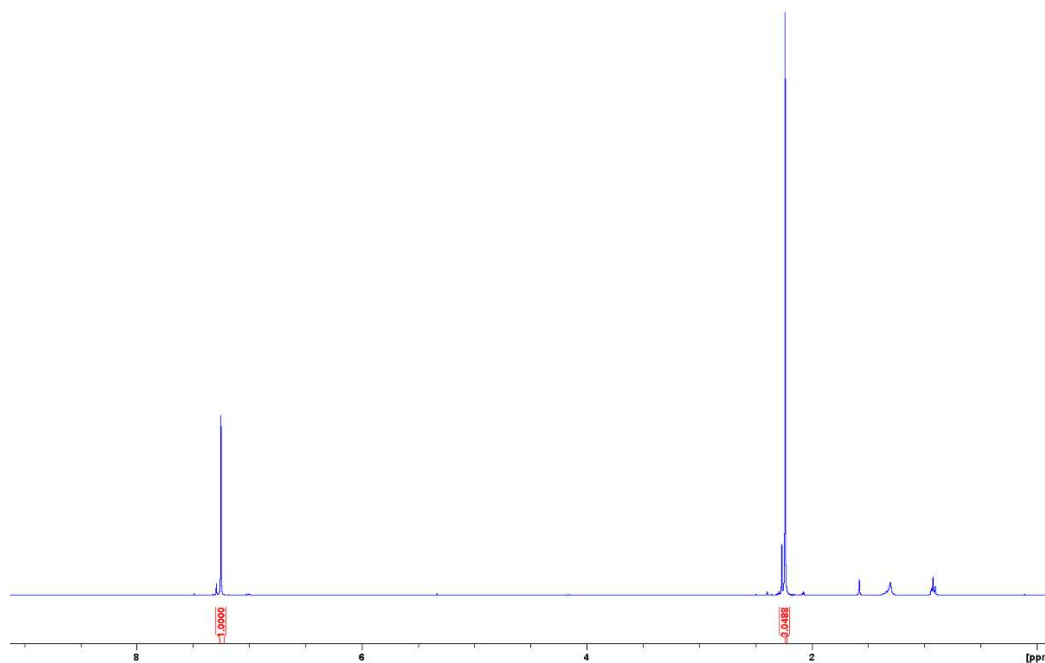
To a stirred solution of 2,3,5-tribromo-4-methylthiophene (6 g, 17.9 mmol) in dry Et₂O (72 mL, 0.25 M) under nitrogen at -78 °C was added *s*-BuLi (1.3 M in cyclohexane, 13.8 mL, 17.9 mmol) dropwise. The resultant mixture was stirred at -78 °C for 10 min, quenched with MeOH (28 mL) at -78 °C, stirred at r.t. for 5 min, concentrated under reduced pressure, diluted in Et₂O (150 mL), and washed with sat. aq. NH₄Cl (70 mL) and then water (70 mL). The organic phase was dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (0-5% EtOAc in hexane) to yield **11** as a colourless oil (4.1 g, 90%).

¹H NMR (400 MHz, CDCl₃) δ 7.25 (s, 1H, H-5), 2.23 (s, 3H, H-6).

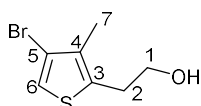
¹³C NMR (100 MHz, CDCl₃) δ 136.4 (C-3), 122.4 (C-5), 112.0 (C-4), 109.1 (C-2), 15.4 (C-6).

Analytical data are consistent with those previously reported.^{3,4}

¹H NMR in CDCl₃:



2-(4-Bromo-3-methylthiophen-2-yl)ethan-1-ol **12**



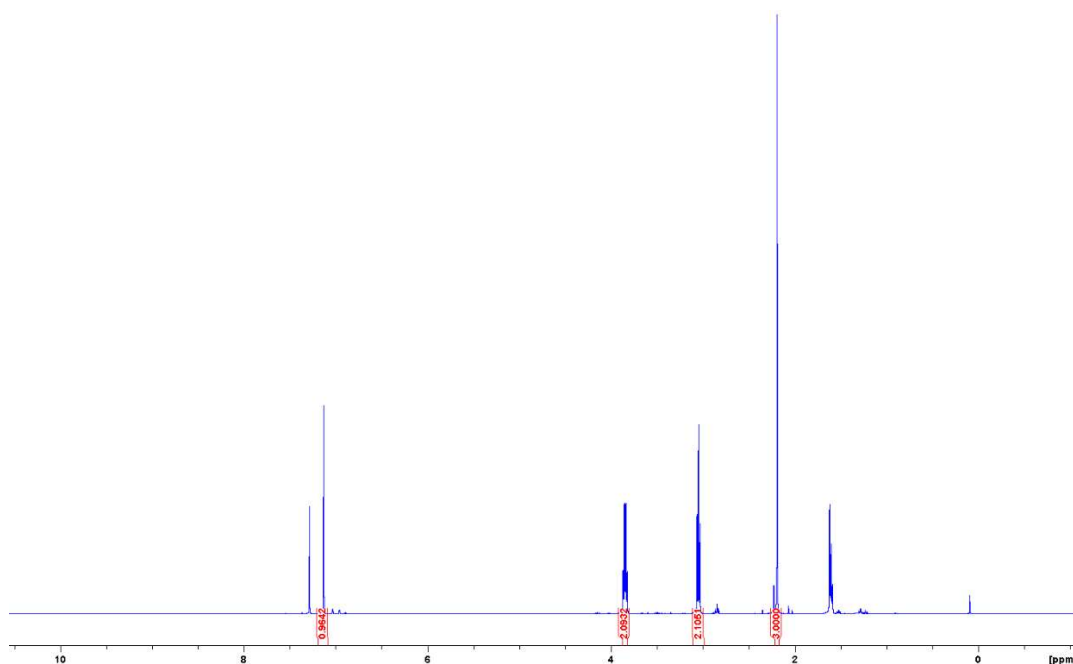
To a stirred solution of **11** (5.6 g, 21.8 mmol) in dry Et₂O (110 mL, 0.2 M) under nitrogen at -78 °C was added *n*-BuLi (1.6 M in hexane, 13.7 mL, 21.8 mmol) dropwise. The resultant mixture was stirred at -78 °C for 15 min, treated with ethylene oxide (2.5-3.3 M in THF, 9.6 mL, 24 mmol) dropwise, stirred at -78 °C for 10 min, treated with boron trifluoride diethyl etherate (2.7 mL, 21.8 mmol) dropwise, and warmed to and stirred at r.t. for 2 h. The reaction mixture was quenched with MeOH (45 mL), concentrated under reduced pressure, diluted in Et₂O (100 mL), and washed with sat. aq. NaHCO₃ (50 mL) and then water (50 mL). The organic phase was dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (30% EtOAc in hexane) to yield **12** as a pale-yellow oil (3.4 g, 71%).

¹H NMR (400 MHz, CDCl₃) δ 7.11 (s, 1H, H-6), 3.83 (t, 2H, *J* = 6.4 Hz, H-1), 3.03 (t, 2H, *J* = 6.4 Hz, H-2), 2.18 (s, 3H, H-7).

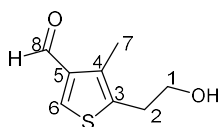
¹³C NMR (100 MHz, CDCl₃) δ 134.0 (C-3), 133.3 (C-4), 119.6 (C-6), 113.3 (C-5), 62.9 (C-1), 32.3 (C-2), 13.7 (C-7).

Analytical data are consistent with those previously reported.^{3,4}

¹H NMR in CDCl₃:



5-(2-Hydroxyethyl)-4-methylthiophene-3-carbaldehyde **13**



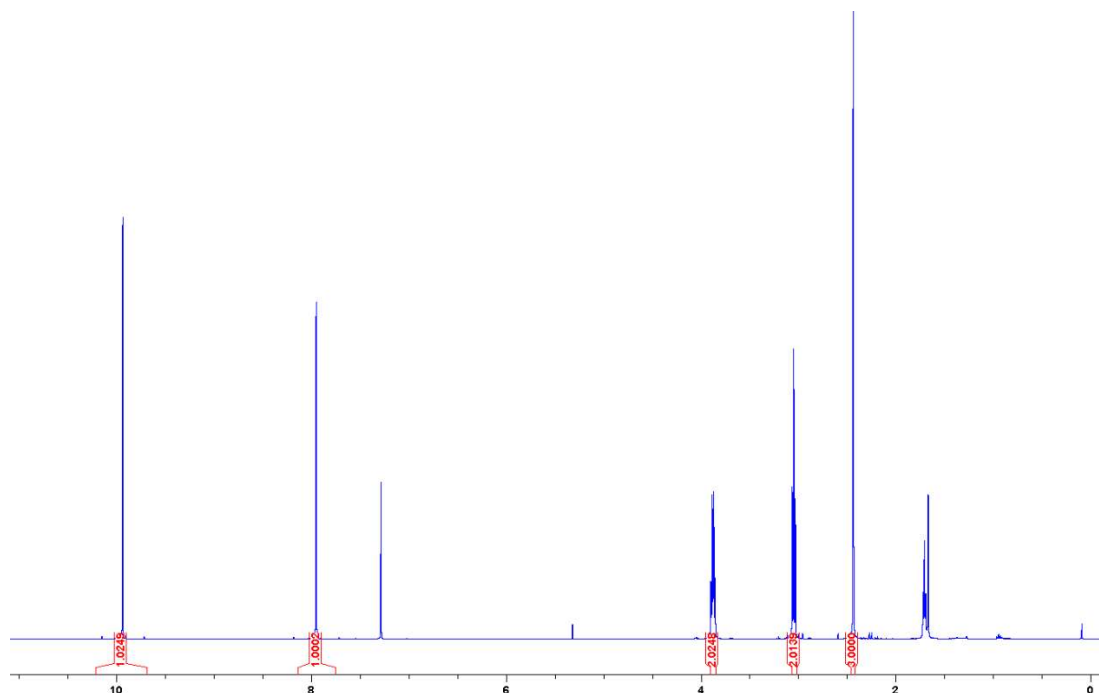
To a stirred solution of **12** (2.2 g, 10 mmol) in dry THF (100 mL, 0.1 M) under nitrogen at 0 °C was added *i*-PrMgCl (2 M in THF, 6 mL, 12 mmol) dropwise. The resultant mixture was stirred at 0 °C for 10 min and then at -78 °C for 5 min, treated with *n*-BuLi (1.6 M in hexane, 13.8 mL, 22 mmol) at -78 °C dropwise, and stirred at -78 °C for 10 min and then at -40 °C for 1 h. The reaction mixture was cooled to -78 °C for 10 min, treated with dry DMF (12.2 mL, 150 mmol) at -78 °C dropwise, stirred at r.t. overnight, quenched with sat. aq. NH₄Cl (50 mL), and extracted with Et₂O (2 x 100 mL). The combined organic phases were washed with water, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (50% EtOAc in hexane) to yield **13** as a colourless oil (1.28 g, 75%).

¹H NMR (400 MHz, CDCl₃) δ 9.93 (s, 1H, H-8), 7.94 (s, 1H, H-6), 3.87 (t, 2H, *J* = 6.4 Hz, H-1), 3.04 (t, 2H, *J* = 6.4 Hz, H-2), 2.42 (s, 3H, H-7).

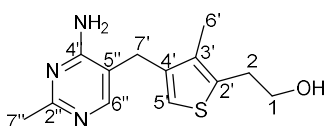
¹³C NMR (100 MHz, CDCl₃) δ 186.2 (C-8), 141.4 (C-5), 137.2 (C-3), 137.1 (C-6), 133.6 (C-4), 62.9 (C-1), 30.9 (C-2), 12.8 (C-7).

Analytical data are consistent with those previously reported.^{3,4}

¹H NMR in CDCl₃:



Deazathiamine 1a



To a stirred solution of **13** (1.7 g, 10 mmol) and β -anilinopropionitrile (1.9 g, 13 mmol) in dry DMSO (33 mL, 0.3 M) under nitrogen at 40 °C was added a freshly prepared solution of NaOMe (2.5 M in dry MeOH, 6 mL, 15 mmol) dropwise. The resultant mixture was stirred at 45 °C overnight. The reaction mixture was quenched with ice water (100 mL) at 0 °C and left in the fridge overnight. The resultant precipitate was filtered under reduced pressure and dried in vacuum to yield the product as a brown solid, which was used in the next step without further purification. To a stirred solution of the resultant solid product and acetamidine hydrochloride (2.08 g, 22 mmol) in dry EtOH (20 mL, 0.5 M) under nitrogen was added a freshly prepared solution of NaOEt (1 M in EtOH, 42 mL, 42 mmol) dropwise. The reaction mixture was stirred at 85 °C for 3 days and then concentrated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1a** as a brown solid (1.7 g, 65% yield over two steps).

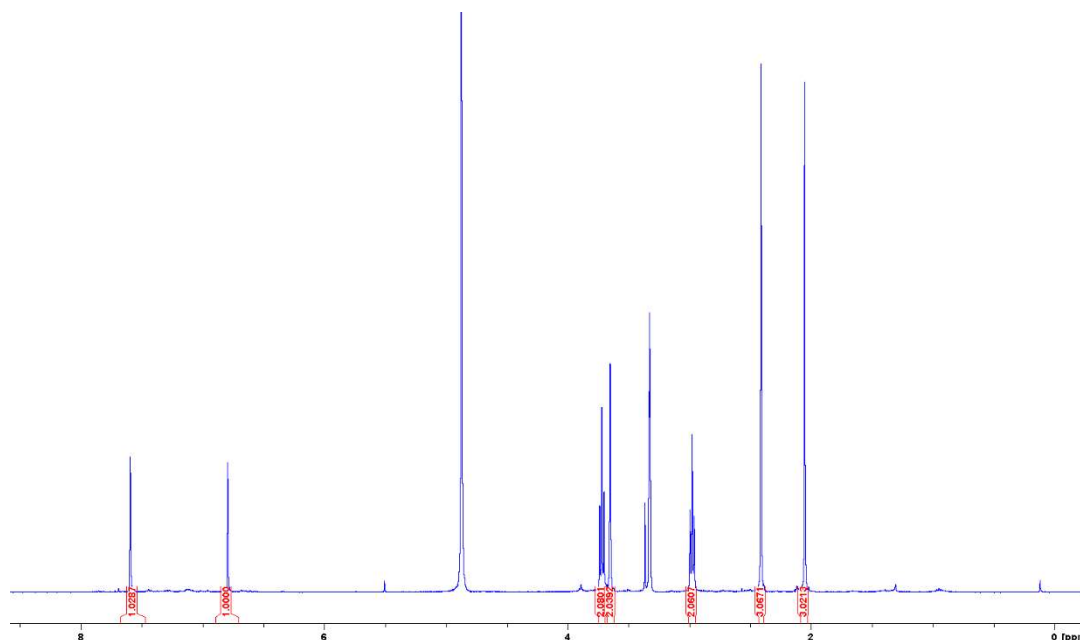
¹H NMR (400 MHz, CD₃OD) δ 7.59 (s, 1H, H-6''), 6.80 (s, 1H, H-5'), 3.72 (t, 2H, J = 6.9 Hz, H-1), 3.66 (s, 2H, H-7'), 2.98 (t, 2H, J = 6.9 Hz, H-2), 2.42 (s, 3H, H-7''), 2.06 (s, 3H, H-6').

¹³C NMR (100 MHz, CD₃OD) δ 164.3 (C-2''), 162.5 (C-4''), 151.6 (C-6''), 136.8 (C-4'), 135.5 (C-3'), 132.6 (C-2'), 118.8 (C-5'), 113.0 (C-5''), 62.3 (C-1), 31.3 (C-2), 27.4 (C-7'), 23.0 (C-7''), 10.9 (C-6').

HRMS (ESI) m/z : [M+H⁺] calculated for C₁₃H₁₇N₃OS: 264.1170; found: 264.1177.

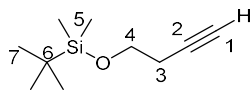
Analytical data are consistent with those previously reported.^{1,3,4}

¹H NMR in CD₃OD:



Preparation of furan-thiamine **1b**:

(*But-3-yn-1-yloxy*)(*tert-butyl*)dimethylsilane **14**



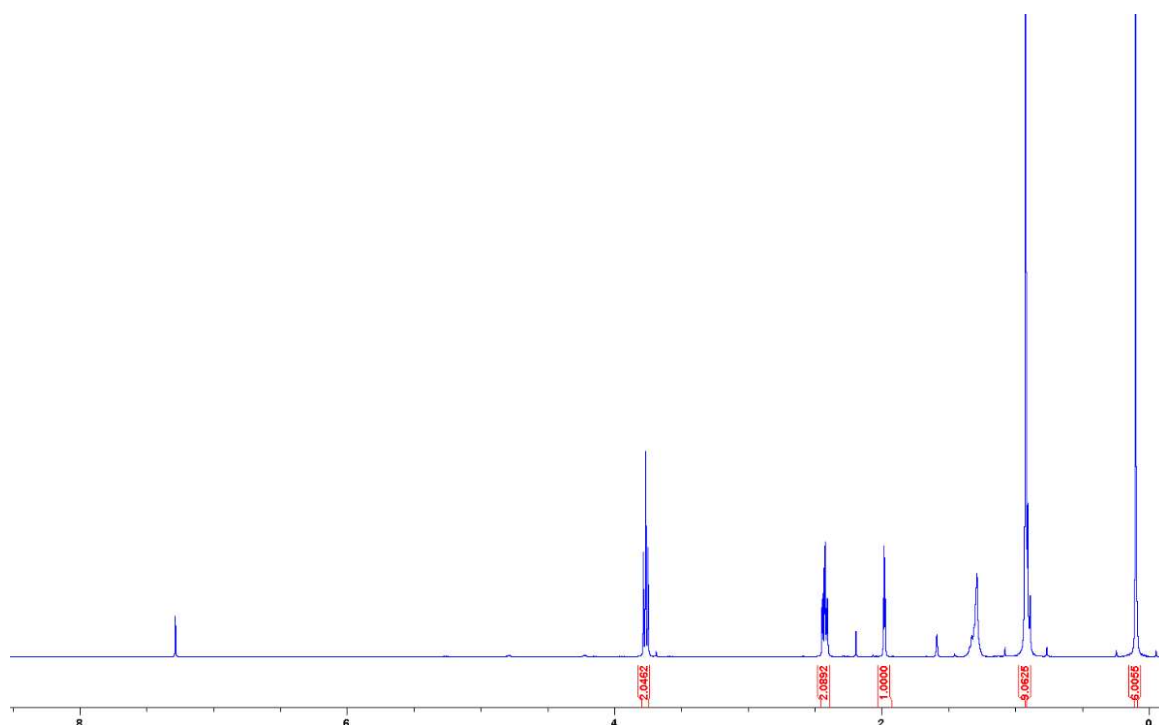
To a stirred solution of 3-butyn-1-ol (10 mL, 132 mmol) in dry DCM (265 mL, 0.5 M) under nitrogen was added imidazole (27 g, 0.4 mol) and TBDMS-Cl (30 g, 0.2 mol). The resultant mixture was stirred at r.t. overnight, concentrated under reduced pressure, diluted in Et₂O (300 mL), washed with aqueous phosphate buffer (pH 7) (3 x 150 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5-10% Et₂O in hexane) to yield **14** as a colourless oil (22.4 g, 92%).

¹H NMR (400 MHz, CDCl₃) δ 3.73 (t, 2H, *J* = 7.2 Hz, H-4), 2.39 (dt, 2H, *J* = 2.4 and 7.2 Hz, H-3), 1.94 (t, 1H, *J* = 2.4 Hz, H-1), 0.89 (s, 9H, H-7), 0.06 (s, 6H, H-5).

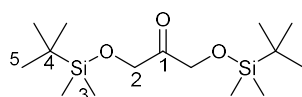
¹³C NMR (100 MHz, CDCl₃) δ 81.5 (C-2), 69.2 (C-1), 61.7 (C-4), 25.8 (C-7), 22.8 (C-6), 18.3 (C-3), -5.3 (C-5).

Analytical data are consistent with those previously reported.^{7,8}

¹H NMR in CDCl₃:



1,2-Bis(*t*-butyldimethylsilyloxy)propan-2-one **15**



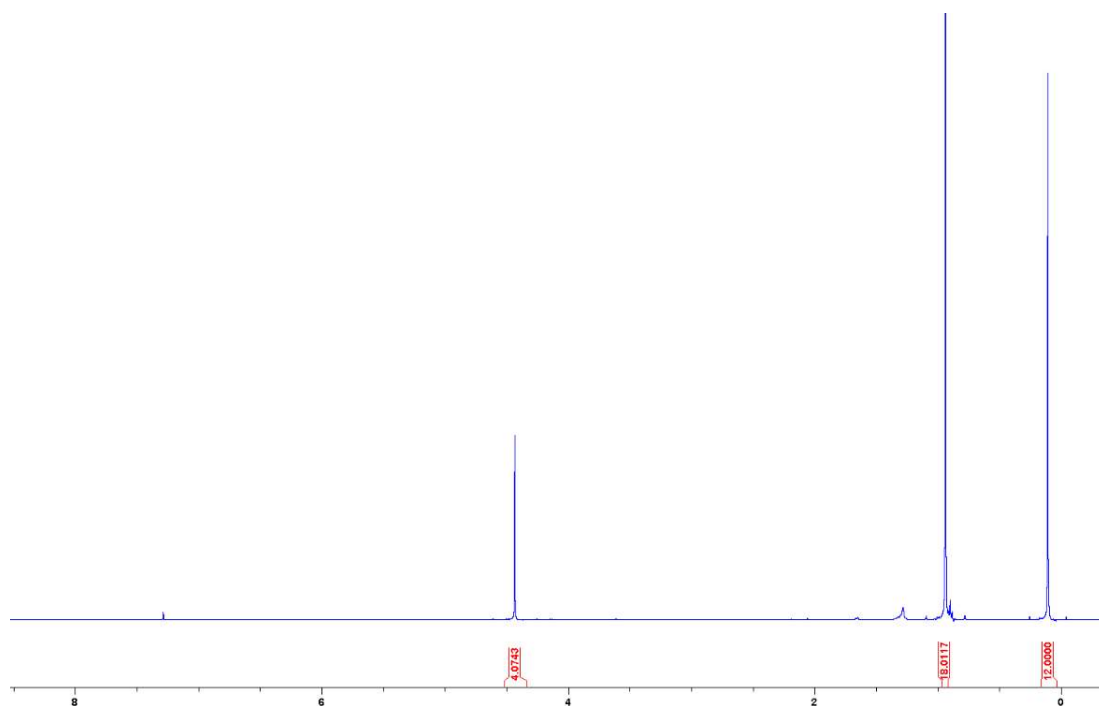
To a stirred solution of dihydroxyacetone (10 g, 0.11 mol) in dry DMF (110 mL, 1 M) under nitrogen was added imidazole (23 g, 0.33 mol) and TBDMS-Cl (42 g, 0.28 mol). The resultant mixture was stirred at r.t. for 3 days, diluted with water (300 mL) and extracted with hexane extensively (5 x 200 mL). The combined organic phases were washed with aqueous phosphate buffer (pH 7) (200 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% EtOAc in hexane) to yield **15** as a colourless oil (30 g, 85%).

¹H NMR (400 MHz, CDCl₃) δ 4.39 (s, 4H, H-2), 0.90 (s, 18H, H-5), 0.07 (s, 12H, H-3).

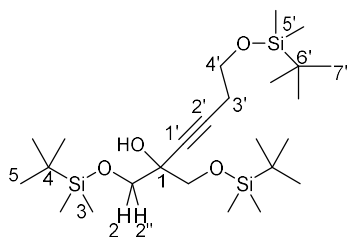
¹³C NMR (100 MHz, CDCl₃) δ 208.6 (C-1), 67.9 (C-2), 25.7 (C-5), 18.3 (C-4), -5.6 (C-3).

Analytical data are consistent with those previously reported.^{7,8}

¹H NMR in CDCl₃:



1,6-Bis-(*t*-butyldimethylsilyloxy)-2-(*t*-butyldimethylsilyloxymethyl)hex-3-yn-2-ol **16**



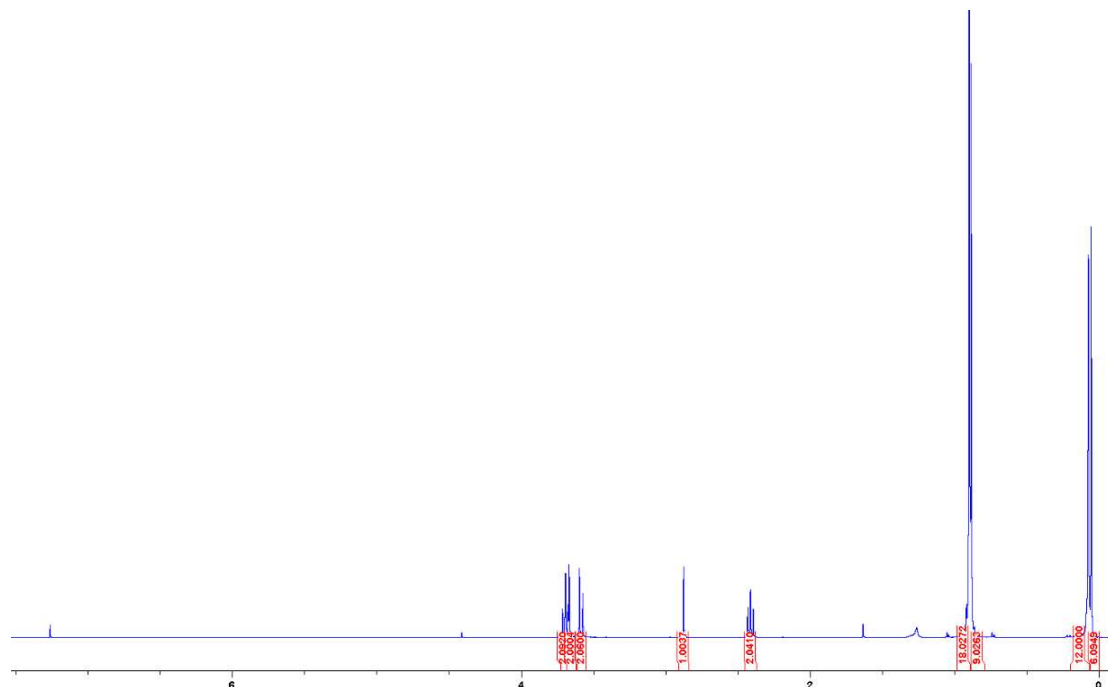
To a stirred solution of **14** (1.84 g, 10 mmol) in dry THF (5 mL, 2 M) under nitrogen at -78°C was added *n*-BuLi (1.6 M in hexane, 6.25 mL, 10 mmol) dropwise. The resultant mixture was stirred at -78°C for 1.5 h, treated with **15** (2 M in dry THF, 5 mL, 10 mmol) slowly at -78°C , and stirred at -78°C for 2 h and then at r.t. for 3 h. The reaction mixture was quenched with sat. aq. NH_4Cl (10 mL) at 0°C , concentrated under reduced pressure (to remove THF), and extracted with Et_2O (100 mL). The organic phase was washed with sat. aq. NaHCO_3 (50 mL), dried over MgSO_4 , filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% EtOAc in hexane) to yield **16** as a colourless oil (4.4 g, 88%).

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 3.68 (t, 2H, $J = 7.6$ Hz, H-4'), 3.67 (d, 2H, $J = 9.5$ Hz, H-2), 3.57 (d, 2H, $J = 9.5$ Hz, H-2''), 2.90 (br, OH), 2.40 (t, 2H, $J = 7.6$ Hz, H-3'), 0.92 (s, 18H, H-5), 0.91 (s, 9H, H-7'), 0.10 (s, 12H, H-3), 0.08 (s, 6H, H-5').

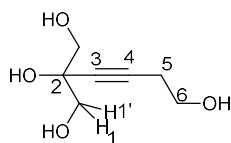
$^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 82.0 (C-1'), 81.1 (C-2'), 71.1 (C-1), 65.8 (C-2), 61.9 (C-4'), 29.0 (C-4 and C-6'), 25.8 (C-5 and C-7'), 23.0 (C-3'), -5.4 (C-3 and C-5').

Analytical data are consistent with those previously reported.^{7,8}

$^1\text{H NMR}$ in CDCl_3 :



2-(Hydroxymethyl)hex-3-yne-1,2,6-triol **17**



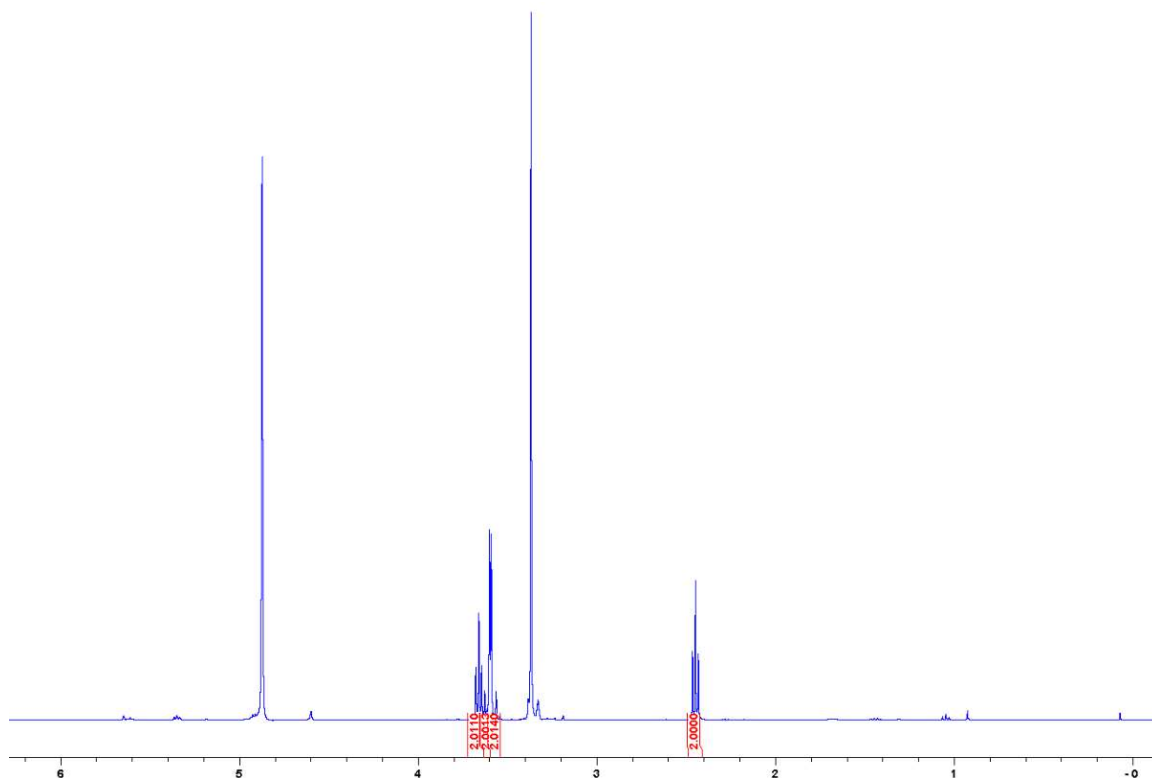
To a stirred solution of **16** (1 g, 2 mmol) in dry THF (2 mL, 1 M) under nitrogen at 0 °C was added TBAF (1 M in THF, 6.6 mL, 6.6 mmol) dropwise. The resultant mixture was stirred at r.t. for 3 h, and treated with CaCO₃ (1.3 g, 200 mg / 1 mmol of TBAF), Dowex 50WX8-400 (4 g, 600 mg / 1 mmol of TBAF) and MeOH (5 mL). The reaction mixture was stirred at r.t. for 3 h, filtered through Celite, and concentrated under reduced pressure to yield **17** as a white solid (314 mg, 98%), which was used in the next step without further purification.

¹H NMR (400 MHz, CD₃OD) δ 3.65 (t, 2H, *J* = 6.8 Hz, H-6), 3.60 (d, 2H, *J* = 10.8 Hz, H-1), 3.55 (d, 2H, *J* = 10.8 Hz, H-1'), 2.45 (t, 2H, *J* = 6.8 Hz, H-5).

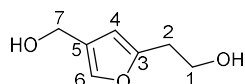
¹³C NMR (100 MHz, CD₃OD) δ 83.1 (C-3), 81.3 (C-4), 71.5 (C-2), 65.9 (C-1), 58.5 (C-6), 19.7 (C-5).

Analytical data are consistent with those previously reported.^{7,8,22}

¹H NMR in CD₃OD:



2-[4-(Hydroxymethyl)furan-2-yl]ethan-1-ol **18**



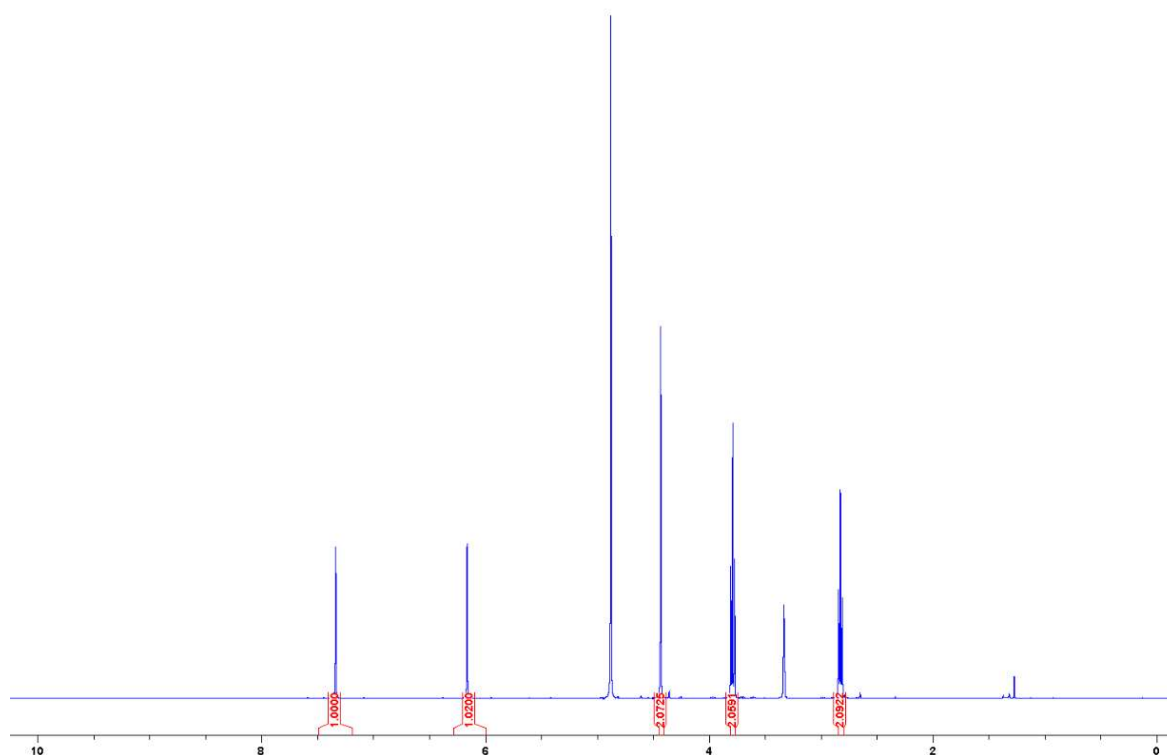
To a stirred suspension of **17** (160 mg, 1 mmol) in warm MeCN (5 mL, 0.2 M) was added Hg(OTf)₂ (50 mg, 0.1 mmol). The resultant mixture was stirred at r.t. for 2 h. The reaction mixture was quenched with TEA (0.1 mL), concentrated under reduced pressure (to remove MeCN), diluted with EtOAc (40 mL), washed with sat. aq. NaHCO₃ (10 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (40% acetone in hexane) to yield **18** as a viscous pale-yellow oil (132 mg, 93%).

¹H NMR (400 MHz, CD₃OD) δ 7.31 (s, 1H, H-6), 6.13 (s, 1H, H-4), 4.40 (s, 2H, H-7), 3.76 (t, 2H, *J* = 6.8 Hz, H-1), 2.79 (t, 2H, *J* = 6.8 Hz, H-2).

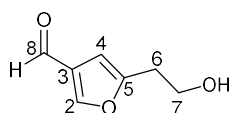
¹³C NMR (100 MHz, CD₃OD) δ 153.8 (C-3), 138.2 (C-6), 126.1 (C-5), 106.2 (C-4), 60.0 (C-1), 55.4 (C-7), 31.2 (C-2).

Analytical data are consistent with those previously reported.^{7,8}

¹H NMR in CD₃OD:



5-(2-Hydroxyethyl)furan-3-carbaldehyde **19**



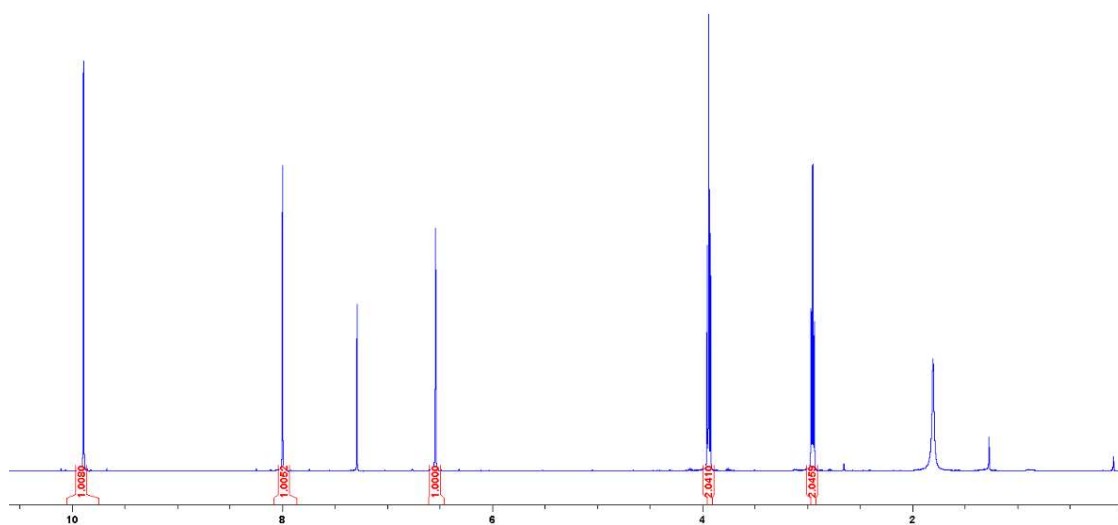
To a stirred suspension of **18** (284 mg, 2 mmol) in chloroform (10 mL, 0.2 M) was added activated manganese oxide (3.5 g, 40 mmol). The resultant mixture was stirred vigorously at r.t. for 2 days, filtered through Celite, and concentrated under reduced pressure. The residue was purified by silica flash chromatography (30% acetone in hexane) to yield **19** as a thin pale-yellow oil (195 mg, 70%).

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 9.85 (s, 1H, H-8), 7.96 (s, 1H, H-2), 6.50 (s, 1H, H-4), 3.90 (t, 2H, $J = 5.2$ Hz, H-7), 2.91 (t, 2H, $J = 5.2$ Hz, H-6).

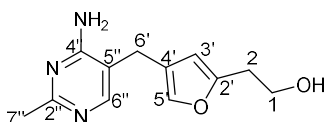
$^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 184.2 (C-8), 156.5 (C-5), 150.5 (C-2), 130.3 (C-3), 103.8 (C-4), 60.6 (C-7), 31.3 (C-6).

Analytical data are consistent with those previously reported.^{7,8}

$^1\text{H NMR}$ in CDCl_3 :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]furan-2-yl}ethan-1-ol (Furan-thiamine) **1b**



To a stirred solution of **19** (0.7 g, 5 mmol) and β -anilinopropionitrile (8.8 g, 6 mmol) in dry DMSO (16.6 mL, 0.3 M) under nitrogen at 40 °C was added a freshly prepared solution of NaOMe (2.5 M in dry MeOH, 3 mL, 7.5 mmol) dropwise. The resultant mixture was stirred at 45 °C overnight, quenched with ice water (60 mL) at 0 °C, and left in the fridge overnight. The resultant precipitate was filtered under reduced pressure and then dried in vacuum to yield the product as a brown solid, which was used in the next step without further purification. To a stirred solution of the resultant solid product and acetamide hydrochloride (0.95 g, 10 mmol) in dry EtOH (10 mL, 0.5 M) under nitrogen was added a freshly prepared solution of NaOEt (1 M in EtOH, 20 mL, 20 mmol) dropwise. The reaction mixture was stirred at 85 °C for 3 days and then concentrated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1b** as a brown solid (0.58 g, 50% yield over two steps).

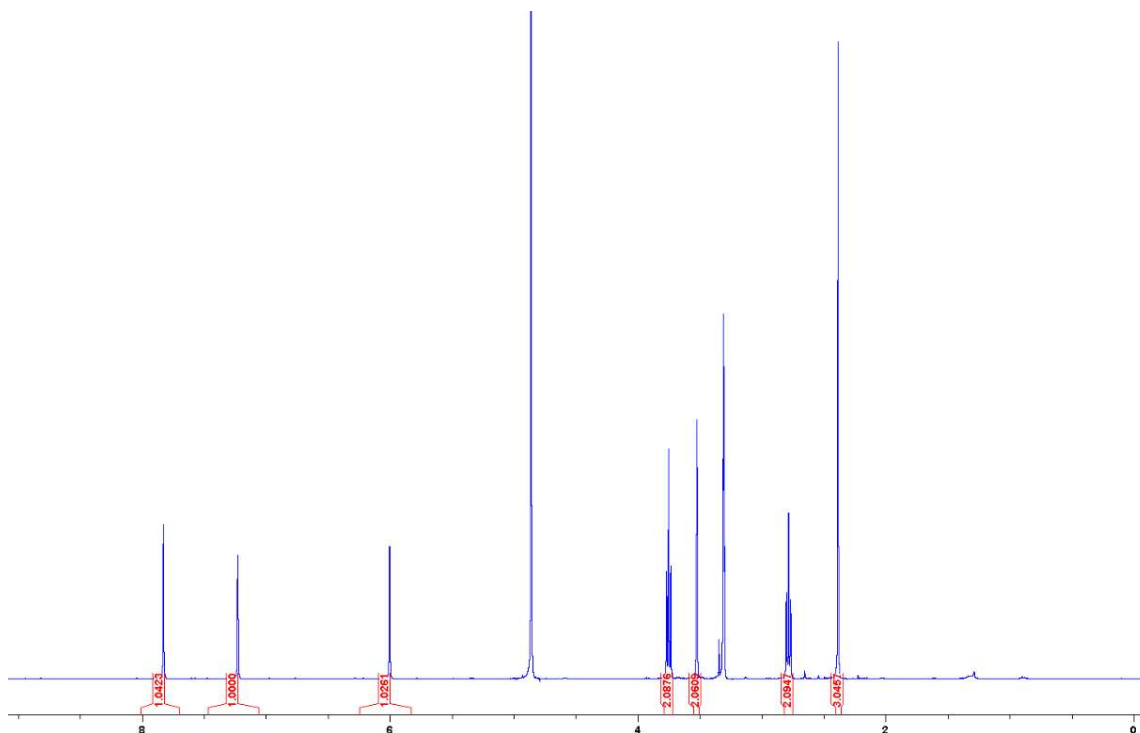
$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.81 (s, 1H, H-6''), 7.24 (s, 1H, H-5'), 6.01 (s, 1H, H-3'), 3.75 (t, 2H, $J = 7.2$ Hz, H-1), 3.53 (s, 2H, H-6'), 2.78 (t, 2H, $J = 7.2$ Hz, H-2), 2.41 (s, 3H, H-7'').

$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 164.9 (C-2''), 162.5 (C-4''), 154.1 (C-2'), 152.9 (C-6''), 138.4 (C-5'), 121.8 (C-4'), 113.3 (C-5''), 107.0 (C-3'), 59.9 (C-1), 31.2 (C-2), 23.4 (C-7''), 22.9 (C-6').

HRMS (ESI) m/z : $[\text{M}+\text{H}^+]$ calculated for $\text{C}_{12}\text{H}_{15}\text{N}_3\text{O}_2$: 234.1242; found: 234.1240.

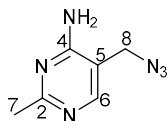
Analytical data are consistent with those previously reported.^{7,8}

$^1\text{H NMR}$ in CD_3OD :



Preparation of triazole-thiamine 1c:

5-(Azidomethyl)-2-methylpyrimidin-4-amine **20**



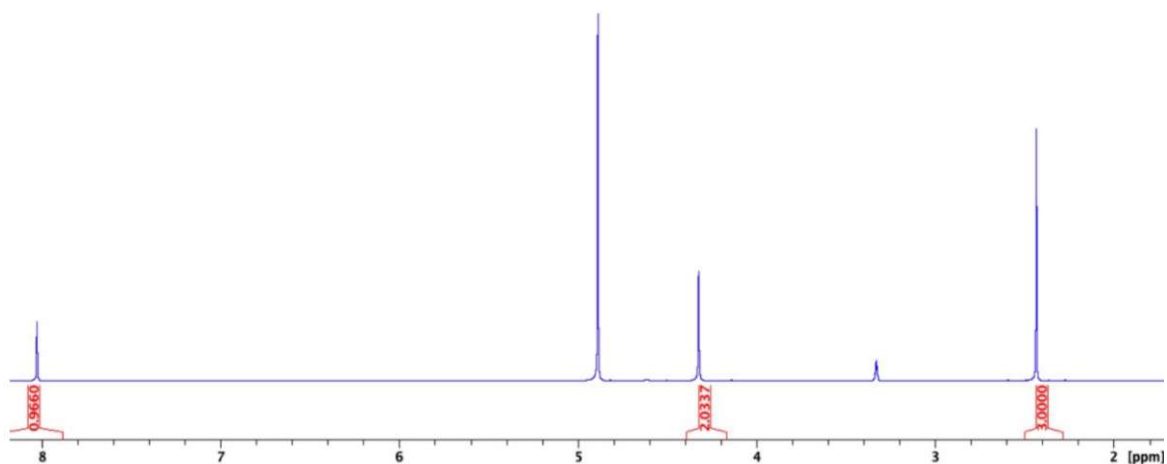
To a stirred solution of thiamine hydrochloride (10 g, 30 mmol) and NaN_3 (5 g, 77 mmol) in water (100 mL, 0.3 M) was added Na_2SO_3 (0.4 g, 3.2 mmol). The resultant mixture was stirred at 70 °C overnight. The reaction mixture was acidified with citric acid to pH 4-5, washed with DCM (200 mL), and basified with potassium carbonate to pH 8-10. Upon product precipitation, the suspension was cooled in an ice bath and filtered under reduced pressure. The filtrate was extracted with EtOAc. The organic phase was dried over MgSO_4 , filtered, and evaporated under reduced pressure. The solid residues were pooled with the precipitate (collected upon filtration) and recrystallised from EtOAc-hexane to yield **20** as a white solid after drying under reduced pressure (3.6 g, 75%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 8.02 (s, 1H, H-6), 4.32 (s, 2H, H-8), 2.43 (s, 3H, H-7).

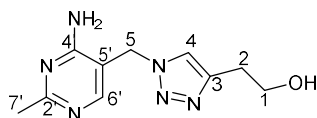
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 167.2 (C-2), 162.3 (C-4), 154.2 (C-6), 108.7 (C-5), 47.8 (C-8), 23.6 (C-7).

Analytical data are consistent with those previously reported.⁹

$^1\text{H NMR}$ in CD_3OD :



2-{1-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-1H-1,2,3-triazol-4-yl}ethan-1-ol
(Triazole-thiamine) **1c**



To a stirred solution of 3-butynol (226 μ l, 3.0 mmol) and **20** (585 mg, 3.6 mmol) in *t*-BuOH (10 mL) and water (5 ml, 0.2 M) was added sodium ascorbate (176 mg, 0.89 mmol) and CuSO₄·5 H₂O (22 mg, 0.09 mmol). The resultant mixture was stirred at r.t. overnight. The reaction mixture was concentrated under reduced pressure, diluted in *n*-BuOH (50 mL), washed with 0.1 M K₂CO₃ (15 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (20% MeOH in DCM) to yield **1c** as a white solid (490 mg, 70%).

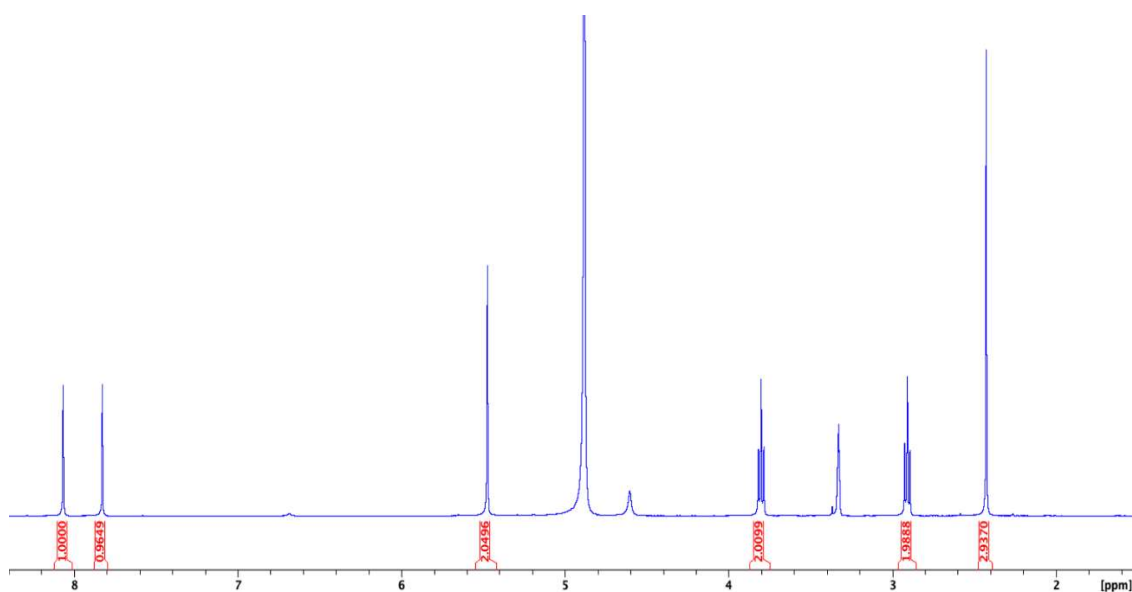
¹H NMR (400 MHz, CD₃OD): δ 8.07 (s, 1H, H-6'), 7.83 (s, 1H, H-4), 5.47 (s, 2H, H-5), 3.80 (t, 2H, *J* = 6.7 Hz, H-1), 2.90 (t, 2H, *J* = 6.7 Hz, H-2), 2.43 (s, 3H, H-7').

¹³C NMR (100 MHz, CD₃SOCD₃) δ 167.8 (C-2'), 162.5 (C-4'), 158.5 (C-6'), 145.1 (C-3), 123.7 (C-4), 109.7 (C-5'), 61.5 (C-1), 47.1 (C-5), 30.3 (C-2), 26.5 (C-7').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₀H₁₄N₆O: 235.1307; found: 235.1311.

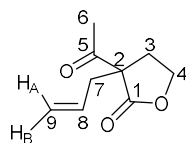
Analytical data are consistent with those previously reported.⁹

¹H NMR in CD₃OD:



Preparation of dethiathiamine 1d:

2-Acetyl-2-allylbutyrolactone **21**



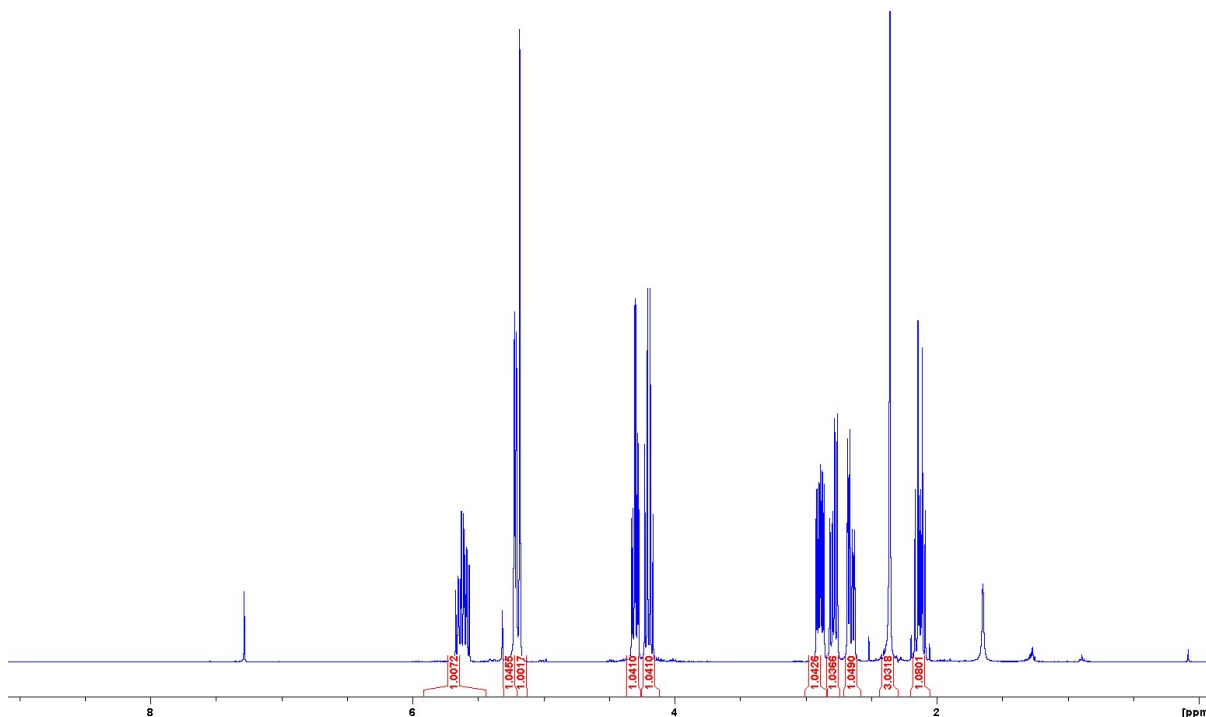
To a stirred solution of 2-acetyl-butylolactone (1.08 mL, 10 mmol) in dry THF (50 mL, 0.2 M) under nitrogen at 0 °C was added sodium hydride (294 mg, 11 mmol) in three portions. The resultant mixture was stirred at r.t. for 30 min, treated with allyl bromide (1.3 mL, 15 mmol) dropwise, and stirred at r.t. overnight. The reaction mixture was concentrated under reduced pressure (to remove THF), diluted with EtOAc (100 mL), washed with aqueous phosphate buffer (pH 7) (30 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (20% EtOAc in Hexane) to yield **21** as a colourless oil (1.46 g, 87%).

¹H NMR (400 MHz, CDCl₃) δ 5.61 (dddd, 1H, *J* = 6.5, 8.0, 11.1 and 16.1 Hz, H-8), 5.21 (m, 1H, H-9_A), 5.19 (m, 1H, H-9_B), 4.30 (td, 1H, *J* = 3.5, 8.9 and 8.9 Hz, H-4_A), 4.22 (td, 1H, *J* = 7.5, 8.9 and 8.9 Hz, H-4_B), 2.90 (ddd, 1H, *J* = 3.5, 7.5 and 13.1 Hz, H-3_A), 2.79 (ddt, 1H, *J* = 0.9, 0.9, 8.0 and 14.5 Hz, H-7_A), 2.66 (ddt, 1H, *J* = 1.4, 1.4, 6.5 and 14.5 Hz, H-7_B), 2.36 (s, 3H, H-6), 2.12 (dt, 1H, *J* = 8.9, 8.9 and 13.1 Hz, H-3_B).

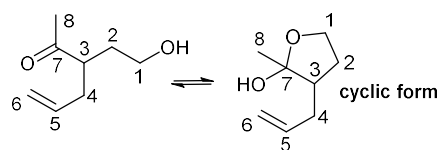
¹³C NMR (100 MHz, CDCl₃) δ 201.8 (C-5), 175.4 (C-1), 131.2 (C-8), 120.4 (C-9), 66.5 (C-4), 61.1 (C-2), 39.2 (C-7), 28.8 (C-3), 25.7 (C-6).

Analytical data are consistent with those previously reported.¹¹

¹H NMR in CDCl₃:



3-Acetylhex-5-enol **22**



To a stirred solution of **21** (168 mg, 1 mmol) in THF (7 mL) and water (3 mL, 0.1 M) was added lithium hydroxide monohydrate (125 mg, 3 mmol). The resultant mixture was stirred at 40 °C for 3 h, concentrated under reduced pressure, diluted with water (10 mL), acidified with conc. HCl to pH = 2-3, and extracted with EtOAc (5 x 20 mL). The combined organic phases were dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (20% EtOAc in Hexane) to yield **22** as a pale-yellow oil (54 mg, 38%).

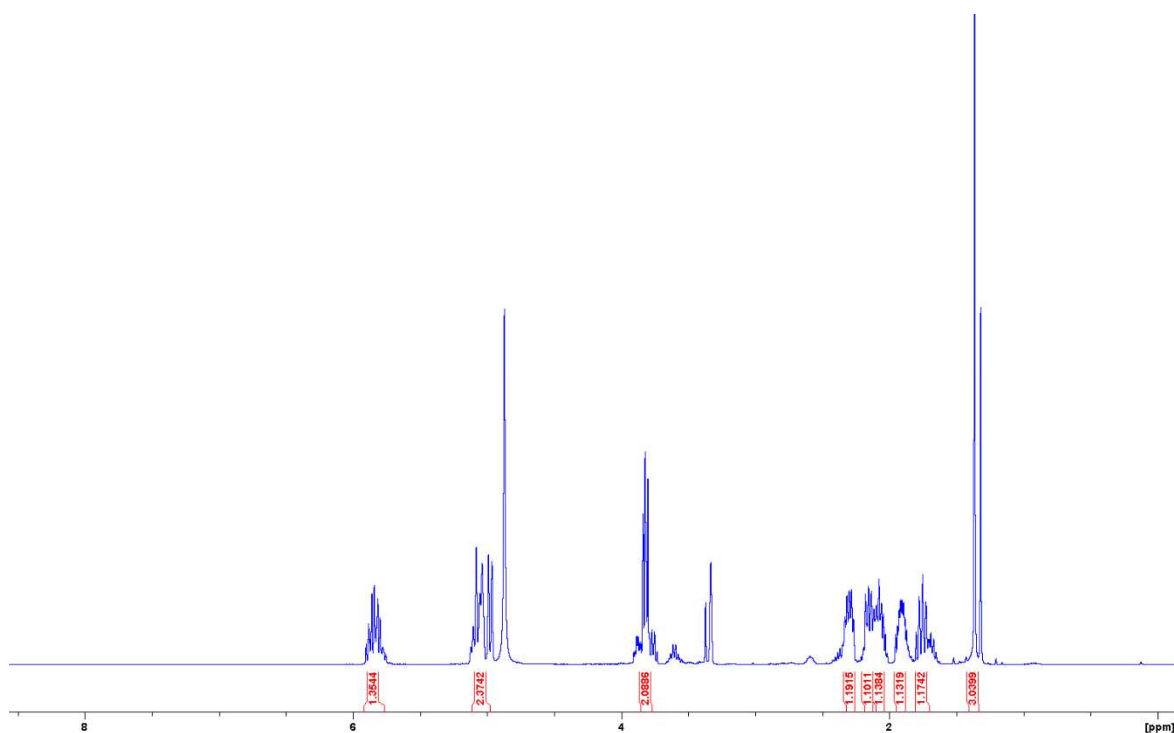
In CDCl₃, both the cyclic and acyclic forms coexist, complicating the analysis of the NMR spectra. In CD₃OD, **22** is mainly in its cyclic form with one diastereoisomer predominating:

¹H NMR (400 MHz, CD₃OD) δ 5.82 (m, 1H, H-5), 4.9-5.1 (m, 2H, H-6), 3.81 (m, 2H, H-1), 2.1-2.3 (m, 2H), 2.05 (m, 1H), 1.90 (m, 1H), 1.75 (m, 1H), 1.37 (s, 3H, H-8).

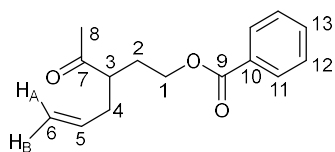
¹³C NMR (100 MHz, CD₃OD) δ 137.3 (C-5), 114.1 (C-6), 106.0 (C-7), 65.1 (C-1), 48.8 (C-3), 33.5 (C-4), 30.1 (C-2), 18.5 (C-8).

Analytical data are consistent with those previously reported.¹¹

¹H NMR in CD₃OD: (the small peaks belong to the protons of the minor diastereomer and of the acyclic form)



3-Acetylhex-5-enyl benzoate **23**



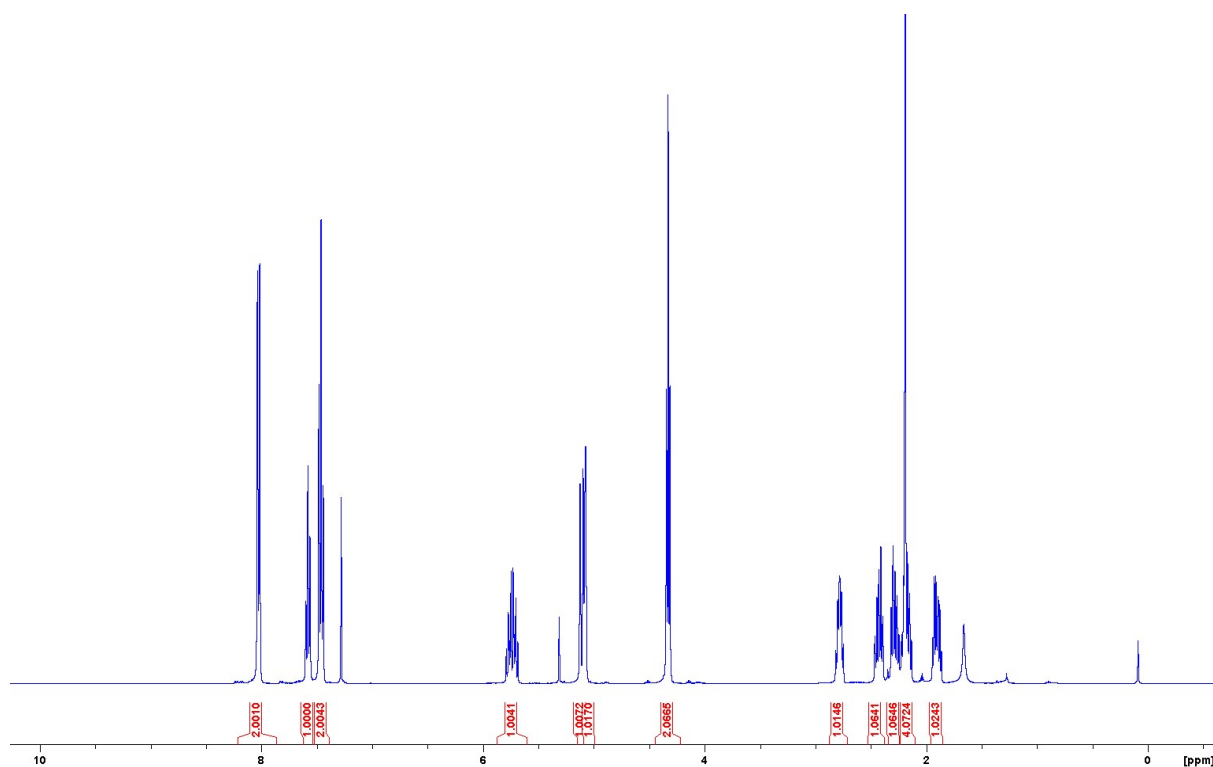
To a stirred solution of **22** (0.71 g, 5 mmol) in dry pyridine (12.5 mL, 0.4 M) under nitrogen at 0 °C was added benzoyl chloride (0.58 mL, 5 mmol) dropwise. The resultant mixture was stirred at 40 °C for 5 h. The reaction mixture was concentrated under reduced pressure, diluted with EtOAc (50 mL), washed with 1 M HCl (50 mL) and then with sat. aq. NaHCO₃ (100 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% EtOAc in Hexane) to yield **23** as a colourless oil (1 g, 83%).

¹H NMR (400 MHz, CDCl₃) δ 8.02 (d, 2H, *J* = 7.7 Hz, H-11), 7.58 (t, 1H, *J* = 7.7 Hz, H-13), 7.47 (t, 2H, *J* = 7.7 Hz, H-12), 5.75 (m, 1H, H-5), 5.10 (m, 1H, H-6_A), 5.09 (m, 1H, H-6_B), 4.32 (t, 2H, *J* = 6.4 Hz, H-1), 2.78 (m, 1H, H-3), 2.41 (m, 1H, H-4_A), 2.30 (m, 1H, H-4_B), 2.21 (s, 3H, H-8), 2.19 (m, 1H, H-2_A), 1.90 (m, 1H, H-2_B).

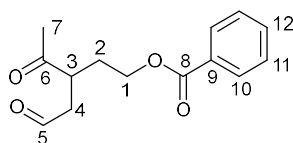
¹³C NMR (100 MHz, CDCl₃) δ 211.1 (C-7), 166.8 (C-9), 134.5 (C-5), 133.1 (C-13), 130.1 (C-10), 129.7 (C-11), 128.4 (C-12), 117.7 (C-6), 63.2 (C-1), 49.3 (C-3), 35.9 (C-4), 29.9 (C-8), 29.5 (C-2).

Analytical data are consistent with those previously reported.¹¹

¹H NMR in CDCl₃:



3-Acetyl-5-oxopentyl benzoate **24**



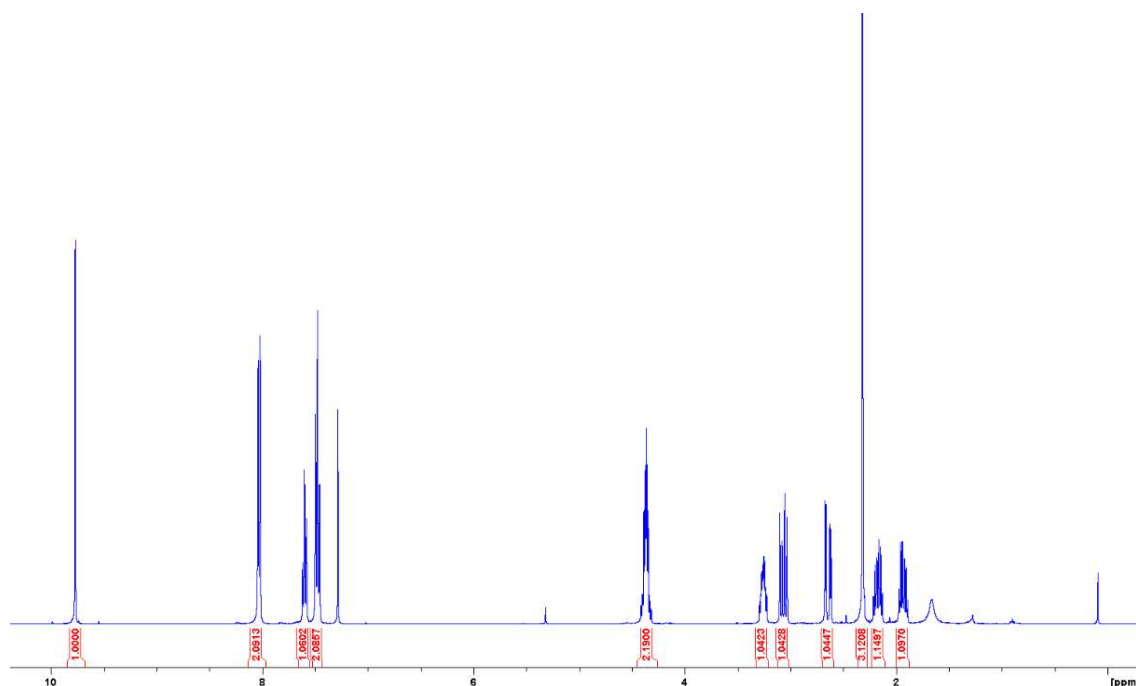
To a stirred solution of **23** (246 mg, 1 mmol) in DCM (5 mL, 0.2 M) at $-78\text{ }^{\circ}\text{C}$ was bubbled ozone from an ozoniser. The resultant mixture was stirred at $-78\text{ }^{\circ}\text{C}$ for 30 min before oxygen was bubbled through the solution to remove the remaining ozone. The reaction mixture was treated with zinc dust (1 g) and glacial acetic acid (0.5 mL), stirred at r.t. for 3 h, filtered, washed with sat. aq. NaHCO_3 (20 mL), dried over MgSO_4 , filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (15% EtOAc in Hexane) to yield **24** as a colourless oil (153 mg, 62%).

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 9.78 (s, 1H, H-5), 8.02 (dd, 2H, $J = 1.4$ and 8.5 Hz, H-10), 7.60 (tt, 1H, $J = 1.4$, 7.4 and 7.4 Hz, H-12), 7.47 (m, 2H, H-11), 4.37 (m, 2H, H-1), 3.27 (dtd, 1H, $J = 4.1$, 6.7, 6.7 and 9.2 Hz, H-3), 3.07 (dd, 1H, $J = 9.2$ and 18.5 Hz, H-4), 2.62 (dd, 1H, $J = 4.1$ and 18.5 Hz, H-4), 2.33 (s, 3H, H-7), 2.18 (dq, 1H, $J = 6.7$, 6.7, 6.7 and 13.0, H-2_A), 1.92 (dq, 1H, $J = 6.7$, 6.7, 6.7 and 13.0, H-2_B).

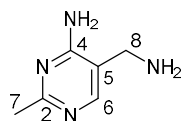
$^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 209.7 (C-6), 200.1 (C-5), 166.7 (C-8), 133.5 (C-12), 129.6 (C-9), 129.5 (C-10), 128.7 (C-11), 62.1 (C-1), 44.9 (C-4), 43.0 (C-3), 30.1 (C-2), 29.7 (C-7).

Analytical data are consistent with those previously reported.¹¹

$^1\text{H NMR}$ in CDCl_3 :



5-(Aminomethyl)-2-methylpyrimidin-4-amine **25**



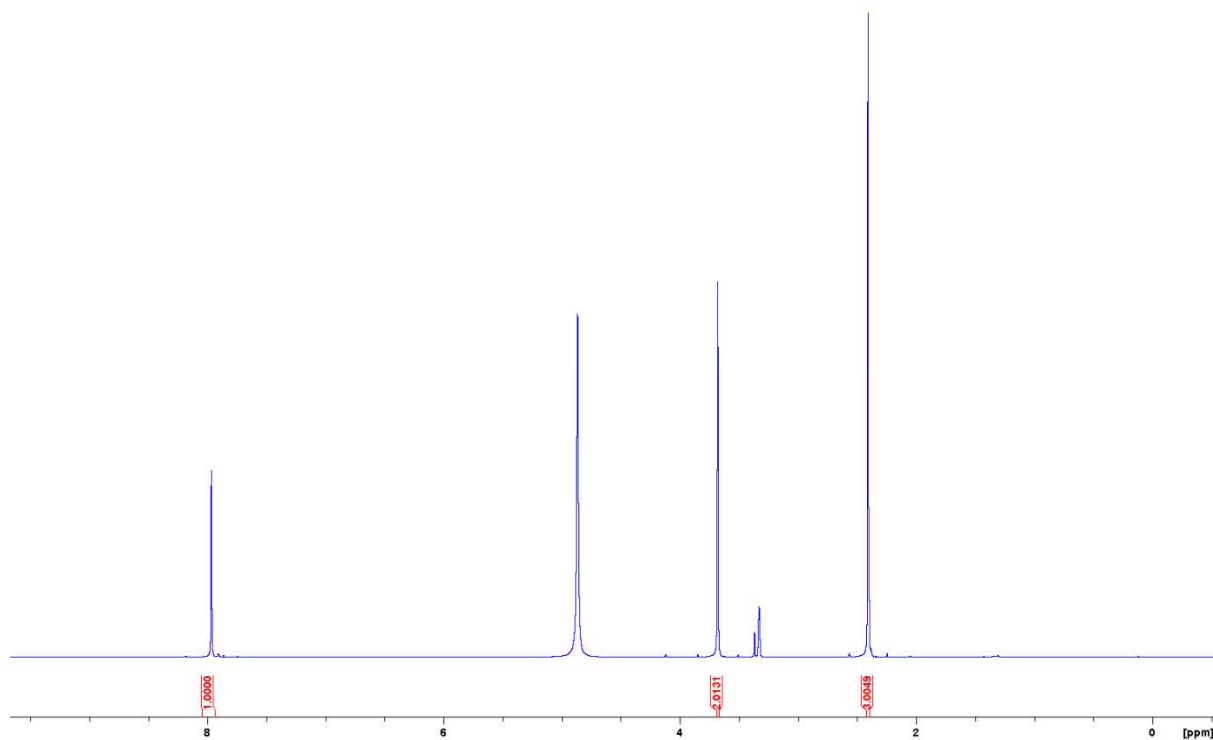
A stirred solution of **20** (488 mg, 3 mmol) in MeOH (15 mL, 0.2 M) at r.t. was treated with 10% Pd/C (50 mg) under nitrogen. The flask was evacuated and flushed with hydrogen gas (three times). The resultant mixture was stirred vigorously at r.t. under an atmosphere of hydrogen (1 atm, H₂ balloon) for 5 h. The reaction mixture was filtered through Celite and then concentrated under reduced pressure to yield **25** as a white solid (412 mg, quantitative), which was used in the next step without further purification.

¹H NMR (400 MHz, CD₃OD) δ 7.97 (s, 1H, H-6), 3.68 (s, 2H, H-8), 2.40 (s, 3H, H-7).

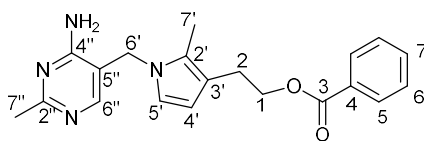
¹³C NMR (100 MHz, CD₃OD) δ 165.5 (C-2), 162.4 (C-4), 152.1 (C-6), 114.8 (C-5), 39.0 (C-8), 23.2 (C-7).

Analytical data are consistent with those previously reported.¹¹

¹H NMR in CD₃OD:



2-(1-((4-Amino-2-methylpyrimidin-5-yl)methyl)-2-methyl-1H-pyrrol-3-yl)ethyl benzoate **26**



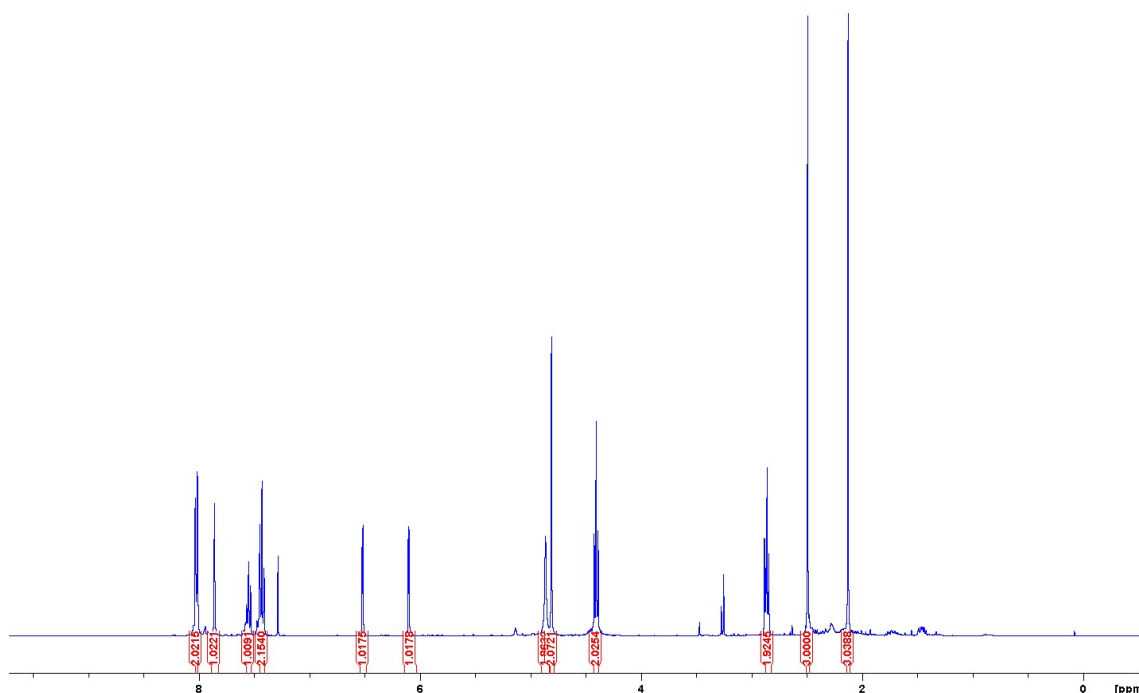
To a stirred solution of **24** (322 mg, 1.3 mmol) in dry DMF (6.5 mL, 0.2 M) under nitrogen was added **25** (270 mg, 2.0 mmol) and some molecular sieves (4 Å). The resultant mixture was stirred at 45 °C for 3 days, filtered, diluted with DCM (300 mL), washed with sat. aq. NaHCO₃ (50 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5-10% MeOH in DCM) to yield product **26** as a white solid (251 mg, 55%).

¹H NMR (400 MHz, CDCl₃) δ 8.02 (m, 2H, H-5), 7.85 (s, 1H, H-6''), 7.55 (m, 1H, H-7), 7.42 (m, 2H, H-6), 6.51 (d, 1H, *J* = 2.8 Hz, H-5'), 6.10 (d, 1H, *J* = 2.8 Hz, H-4'), 4.86 (br, 2H, NH₂), 4.81 (s, 2H, H-6'), 4.41 (t, 2H, *J* = 7.2 Hz, H-1), 2.86 (t, 2H, *J* = 7.2 Hz, H-2), 2.50 (s, 3H, H-7''), 2.13 (s, 3H, H-7').

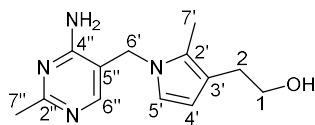
¹³C NMR (100 MHz, CDCl₃) δ 167.5 (C-2''), 166.6 (C-3), 161.2 (C-4''), 155.1 (C-6''), 132.8 (C-7), 130.4 (C-4), 129.5 (C-5), 128.3 (C-6), 125.9 (C-2'), 119.0 (C-5'), 116.9 (C-3'), 109.9 (C-5''), 109.3 (C-4'), 65.8 (C-1), 46.0 (C-6'), 26.2 (C-2), 25.6 (C-7''), 10.0 (C-7').

Analytical data are consistent with those previously reported.¹¹

¹H NMR in CDCl₃:



2-{1-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-2-methyl-1H-pyrrol-3-yl}ethan-1-ol
(Pyrrothiamine) **1d**



To a stirred solution of **26** (70 mg, 0.2 mmol) in dry MeOH (1 mL, 0.2 M) under nitrogen was added potassium carbonate (55 mg, 0.4 mmol). The resultant mixture was stirred at r.t. for 3 h and then concentrated under reduced pressure. The residue was purified by silica flash chromatography (10% MeOH in DCM) to yield product **1d** as a white solid (39 mg, 80%).

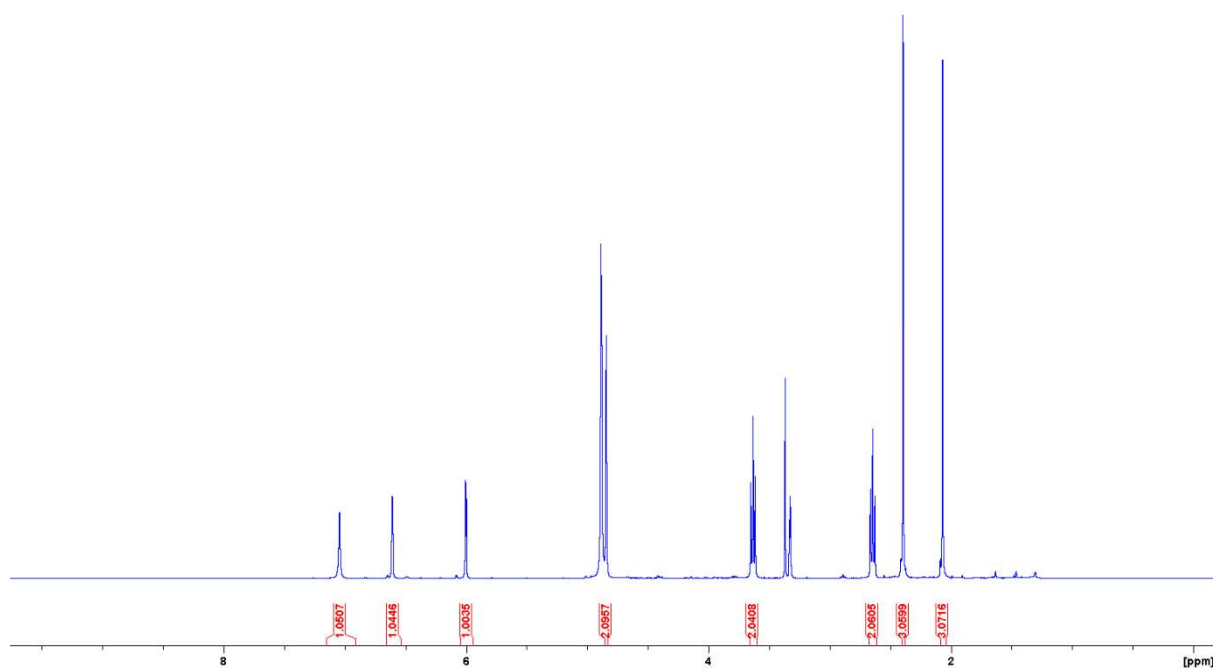
¹H NMR (400 MHz, CD₃OD) δ 7.07 (s, 1H, H-6''), 6.60 (d, 1H, *J* = 2.7 Hz, H-5'), 6.01 (d, 1H, *J* = 2.7 Hz, H-4'), 4.85 (s, 2H, H-6'), 3.64 (t, 2H, *J* = 7.8 Hz, H-1), 2.65 (t, 2H, *J* = 7.8 Hz, H-2), 2.41 (s, 3H, H-7''), 2.07 (s, 3H, H-7').

¹³C NMR (100 MHz, CD₃OD) δ 165.8 (C-2''), 160.6 (C-4''), 151.5 (C-6''), 124.7 (C-2'), 119.2 (C-5'), 116.9 (C-3'), 112.2 (C-5''), 108.1 (C-4'), 63.2 (C-1), 43.5 (C-6'), 29.8 (C-2), 23.5 (C-7''), 8.1 (C-7').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₃H₁₈N₄O: 247.1558; found: 247.1570.

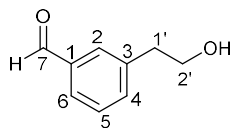
Analytical data are consistent with those previously reported.¹¹

¹H NMR in CD₃OD:



Preparation of *meta*-phenyl-thiamine 1a:

3-(2-Hydroxyethyl)benzaldehyde **27**



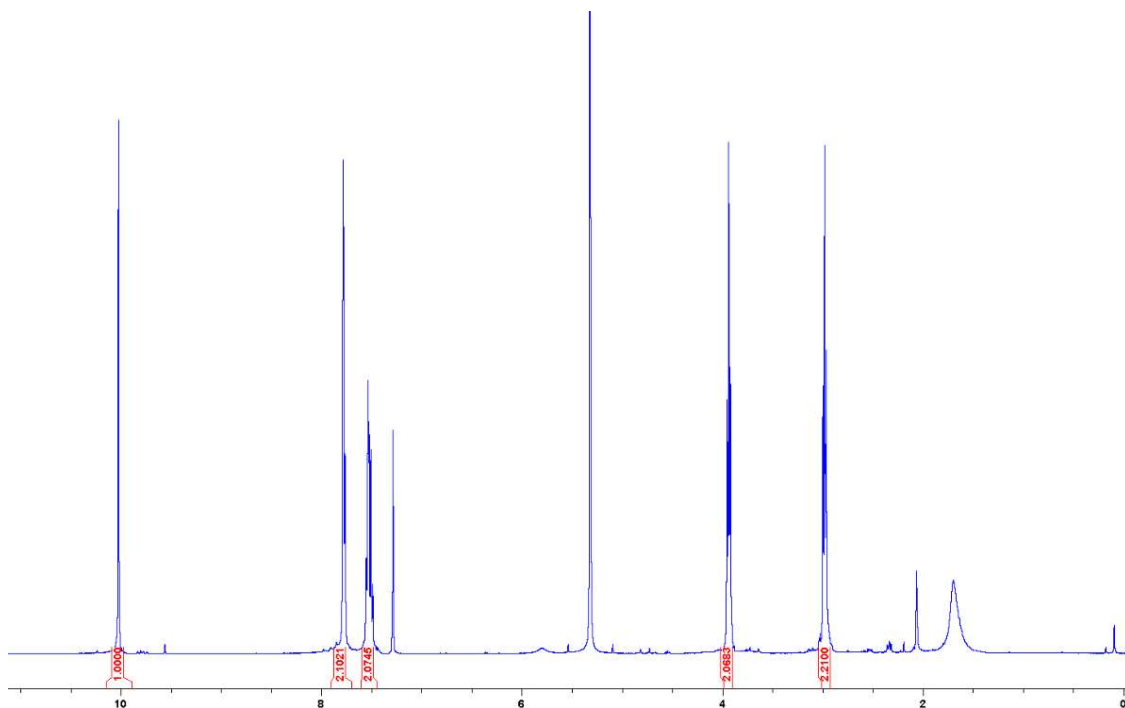
To a stirred solution of 2-(3-bromophenyl)ethanol (2 g, 10 mmol) in dry THF (100 mL, 0.1 M) under nitrogen at 0 °C was added *i*-PrMgCl (2 M in THF, 6 mL, 12 mmol) dropwise. The resultant mixture was stirred at 0 °C for 10 min and then at -78 °C for 5 min, treated with *n*-BuLi (1.6 M in hexane, 13.8 mL, 22 mmol) at -78 °C dropwise, and stirred at -78 °C for 10 min and then at -40 °C for 1 h. The reaction mixture was cooled to -78 °C for 10 min, treated with dry DMF (12.2 mL, 150 mmol) at -78 °C dropwise, stirred at r.t. overnight, quenched with sat. aq. NH₄Cl (50 mL), and extracted with Et₂O (5 x 150 mL). The combined organic phases were washed with water (200 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (50% Et₂O in hexane) to yield **27** as a colourless oil (1.17 g, 78%).

¹H NMR (400 MHz, CDCl₃) δ 10.02 (s, 1H, H-7), 7.74-7.78 (m, 2H, H-2 and H-6), 7.48-7.57 (m, 2H, H-4 and H-5), 3.94 (t, 2H, *J* = 6.6 Hz, H-2'), 2.98 (t, 2H, *J* = 6.6 Hz, H-1').

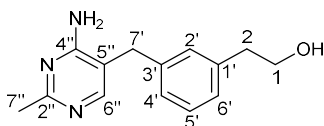
¹³C NMR (100 MHz, CDCl₃) δ 192.2 (C-7), 140.2 (C-1), 136.7 (C-3), 135.4 (C-4), 130.0 (C-6), 129.5 (C-5), 128.5 (C-2), 63.3 (C-2'), 38.9 (C-1').

Analytical data are consistent with those previously reported.⁴

¹H NMR in CDCl₃:



2-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl}ethan-1-ol (Meta-phenyl-thiamine) **1e**



To a stirred solution of **27** (1.17 g, 7.8 mmol) and β -anilinopropionitrile (1.5 g, 10.3 mmol) in dry DMSO (26 mL, 0.3 M) under nitrogen at 40 °C was added a freshly prepared solution of NaOMe (2.5 M in dry MeOH, 5 mL, 12.5 mmol) dropwise. The resultant mixture was stirred at 45 °C overnight, quenched with ice water (80 mL) at 0 °C, and left in the fridge overnight. The resultant precipitate was filtered under reduced pressure and dried in vacuum to yield the product as a brown solid, which was used in the next step without further purification. To a stirred solution of the resultant solid product and acetamidine hydrochloride (1.28 g, 17.2 mmol) in dry EtOH (15.6 mL, 0.5 M) under nitrogen was added a freshly prepared solution of NaOEt (1 M in EtOH, 33 mL, 33 mmol) dropwise. The reaction mixture was stirred at 85 °C for 3 days and then concentrated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1e** as a brown solid (0.3 g, 16% yield over two steps).

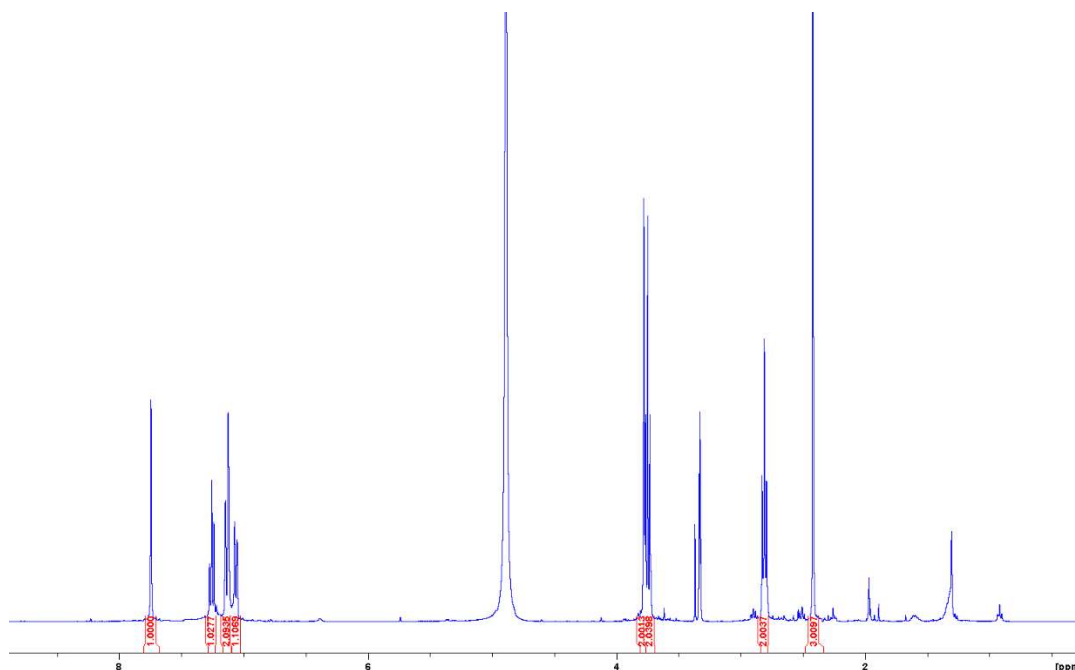
¹H NMR (400 MHz, CD₃OD) δ 7.75 (s, 1H, H-6''), 7.26 (t, 1H, $J = 7.5$ Hz, H-5'), 7.12-7.14 (m, 2H, H-2' and H-4'), 7.06 (d, 1H, $J = 7.5$ Hz, H-6'), 3.78 (s, 2H, H-7), 3.75 (t, 2H, $J = 7.1$ Hz, H-1), 2.81 (t, 2H, $J = 7.1$ Hz, H-2), 2.42 (s, 3H, H-7'').

¹³C NMR (100 MHz, CD₃OD) δ 164.7 (C-2''), 162.3 (C-4''), 152.6 (C-6''), 139.7 (C-3'), 137.5 (C-1'), 129.1 (C-4'), 128.4 (C-6'), 127.1 (C-2'), 126.2 (C-5'), 114.0 (C-5''), 62.8 (C-1), 38.8 (C-2), 32.9 (C-7'), 23.1 (C-7'').

HRMS (ESI) m/z : [M+H⁺] calculated for C₁₄H₁₇N₃O: 244.1449; found: 244.1455.

Analytical data are consistent with those previously reported.⁴

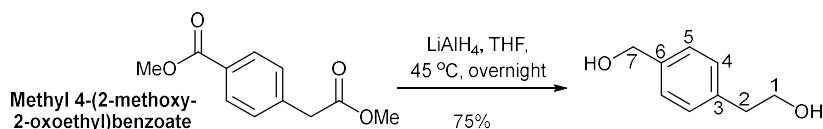
¹H NMR in CD₃OD:



Preparation of *para*-phenyl-thiamine **1f**:

2-[4-(Hydroxymethyl)phenyl]ethan-1-ol

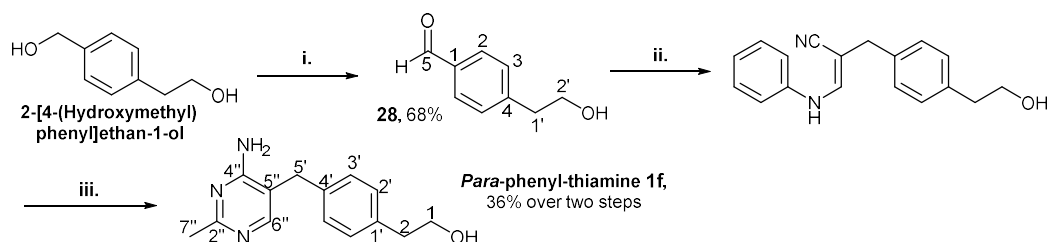
Despite being commercially available, it is cheaper to prepare from:



To a stirred solution of methyl 4-(2-methoxy-2-oxoethyl)benzoate (835 mg, 4 mmol) in dry THF (10 mL, 0.4 M) at 0 °C was added LiAlH₄ (2.4 M in THF, 8.3 mL, 20 mmol) dropwise. The resultant mixture was stirred at 45 °C overnight, quenched with *i*-PrOH (20 mL) at 0 °C slowly, diluted with cold water (20 mL), and extracted with EtOAc (200 mL). The organic phase was dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (30% EtOAc in hexane) to yield 2-[4-(hydroxymethyl)phenyl]ethan-1-ol as a colourless oil (456 mg, 75%).

¹H NMR (400 MHz, CDCl₃) δ 7.33 (d, 2H, *J* = 8.0 Hz, H-5), 7.24 (d, 2H, *J* = 8.0 Hz, H-4), 4.66 (s, 2H, H-7), 3.86 (t, 2H, *J* = 6.6 Hz, H-1), 2.88 (t, 2H, *J* = 6.6 Hz, H-2).

¹³C NMR (100 MHz, CDCl₃) δ 139.1 (C-6), 138.0 (C-3), 129.2 (C-5), 127.4 (C-4), 65.1 (C-7), 63.6 (C-1), 38.9 (C-2).



To a stirred solution of 2-[4-(hydroxymethyl)phenyl]ethan-1-ol (760 mg, 5 mmol) in chloroform (25 mL, 0.2 M) was added activated manganese oxide (8.7 g, 0.1 mol). The reaction mixture was stirred vigorously at r.t. overnight, filtered through Celite, and concentrated under reduced pressure to yield **28** as a thin pale-yellow oil (510 mg, 68%), which was used in the next step without further purification. To a stirred solution of **28** and β-anilinopropionitrile (0.95 g, 6.5 mmol) in dry DMSO (17 mL, 0.3 M) under nitrogen at 40 °C was added a freshly prepared solution of NaOMe (2.5 M in dry MeOH, 3 mL, 7.5 mmol) dropwise. The reaction mixture was stirred at 45 °C overnight, quenched with ice water (60 mL) at 0 °C, and left in the fridge overnight. The precipitate was filtered under reduced pressure and dried in vacuum to yield the product as a brown solid, which was used in the next step without further purification. To a stirred solution of the resultant solid product and acetamidine hydrochloride (0.82 g, 11 mmol) in dry EtOH (10 mL, 0.5 M) under nitrogen was added a freshly prepared solution of NaOEt (1 M in EtOH, 21 mL, 21 mmol) dropwise. The reaction mixture was stirred at 85 °C for 3 days and then concentrated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1f** as a brown solid (297 mg, 36% yield over two steps).

4-(2-Hydroxyethyl)benzaldehyde **28**

¹H NMR (400 MHz, CDCl₃) δ 10.0 (s, 1H, H-5), 7.85 (d, 2H, *J* = 8.0 Hz, H-2), 7.43 (d, 2H, *J* = 8.0 Hz, H-3), 3.94 (t, 2H, *J* = 6.5 Hz, H-2'), 2.98 (t, 2H, *J* = 6.5 Hz, H-1').

¹³C NMR (100 MHz, CDCl₃) δ 192.2 (C-5), 146.6 (C-1), 134.7 (C-4), 130.2 (C-2), 129.5 (C-3), 63.0 (C-2'), 38.9 (C-1').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₉H₁₀O₂: 151.0758; found: 151.0766.

2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl}ethan-1-ol **1f**

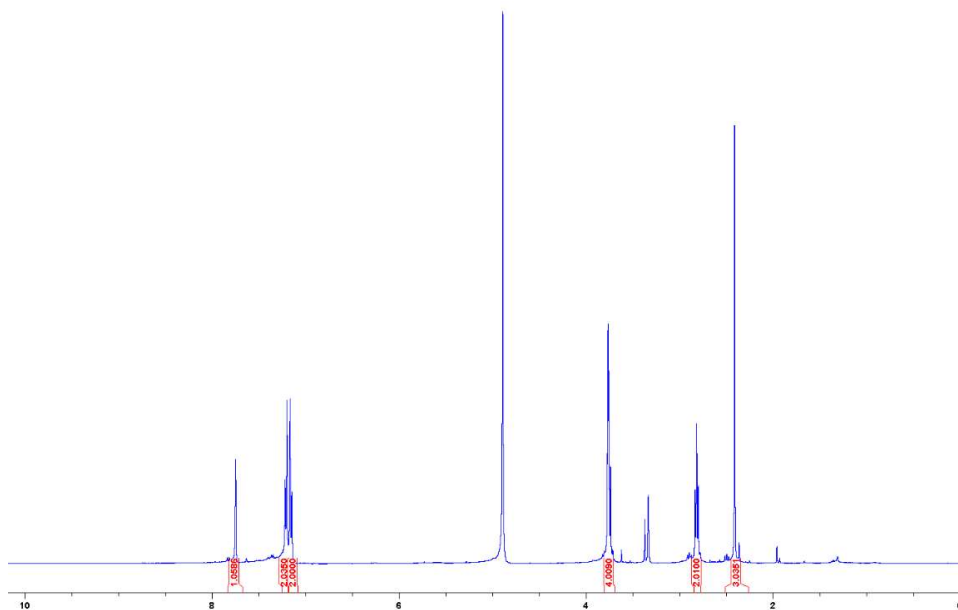
m.p. 177-178 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.75 (s, 1H, H-6''), 7.20 (d, 2H, *J* = 7.9 Hz, H-2'), 7.15 (d, 2H, *J* = 7.9 Hz, H-3'), 3.75 (m, 4H, H-1 and H-5'), 2.81 (t, 2H, *J* = 6.9 Hz, H-2), 2.41 (s, 3H, H-7'').

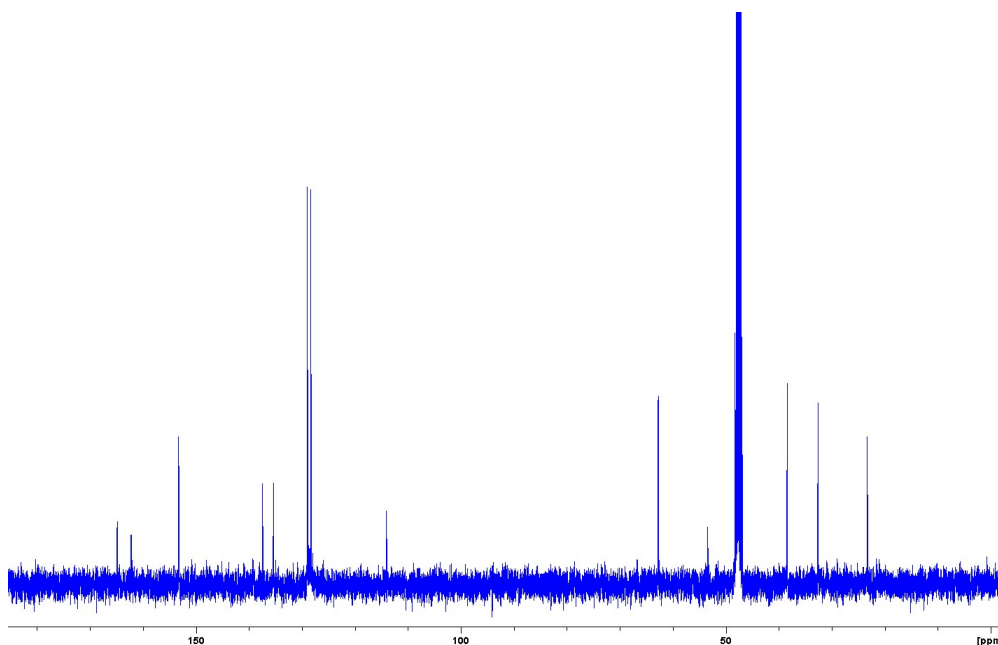
¹³C NMR (100 MHz, CD₃OD) δ 164.9 (C-2''), 162.3 (C-4''), 153.3 (C-6''), 137.5 (C-4'), 135.4 (C-1'), 129.0 (C-3'), 128.3 (C-2'), 114.1 (C-5''), 63.0 (C-1), 38.3 (C-2), 32.5 (C-5'), 23.4 (C-7'').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₄H₁₇N₃O: 244.1449; found: 244.1450.

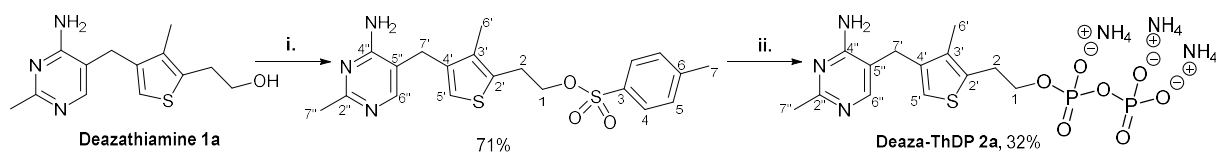
¹H NMR in CD₃OD:



¹³C NMR in CD₃OD:



Preparation of deaza-ThDP **2a**:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl}ethyl 4-methylbenzene-1-sulfonate

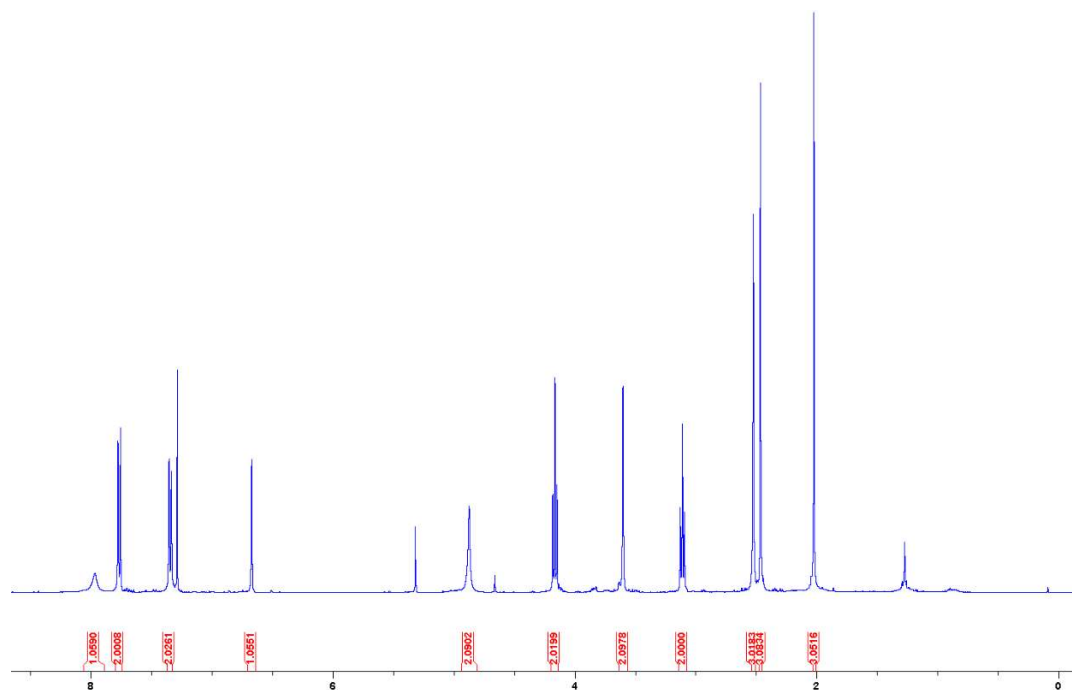
To a stirred solution of **1a** (53 mg, 0.2 mmol) in dry pyridine (1 mL, 0.2 M) under nitrogen at 0 °C was added *p*-TsCl (190 mg, 1 mmol). The resultant mixture was stirred at 0 °C for 4 h, quenched with cold 1 M HCl (1 mL), diluted with water (15 mL), neutralised with sodium bicarbonate to pH 7, and extracted with DCM (50 mL). The organic phase was washed with sat. aq. Cu₂SO₄, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1a**'s tosylate as a brown solid (59 mg, 71%).

¹H NMR (400 MHz, CDCl₃): δ 7.97 (s, 1H, H-6''), 7.77 (d, 2H, *J* = 8.4 Hz, H-4), 7.34 (d, 2H, *J* = 8.4 Hz, H-5), 6.66 (s, 1H, H-5'), 4.88 (br, NH₂), 4.16 (t, 2H, *J* = 7.0 Hz, H-1), 3.60 (s, 2H, H-7'), 3.11 (t, 2H, *J* = 7.0 Hz, H-2), 2.52 (s, 3H, H-7''), 2.47 (s, 3H, H-7), 2.02 (s, 3H, H-6').

¹³C NMR (100 MHz, CDCl₃) δ 166.1 (C-2''), 161.3 (C-4''), 155.5 (C-6''), 144.6 (C-3), 136.7 (C-4'), 133.5 (C-3'), 132.9 (C-6), 132.8 (C-2'), 129.6 (C-5), 127.6 (C-4), 119.4 (C-5'), 111.3 (C-5''), 69.3 (C-1), 28.9 (C-2), 27.9 (C-7'), 25.2 (C-7''), 21.4 (C-7), 12.2 (C-6').

Analytical data are consistent with those previously reported.^{1,3}

¹H NMR of **1a**'s tosylate in CDCl₃:



Deazathiamine diphosphate (Deaza-ThDP) **2a**

To a stirred solution of **1a**'s tosylate (13 mg, 0.031 mmol) in dry MeCN (0.15 mL, 0.2 M) under nitrogen at 0 °C was added tris(tetra-*n*-butylammonium) hydrogen pyrophosphate (56 mg, 0.062 mmol). The resultant mixture was stirred at r.t. for 6 h. The reaction mixture was concentrated under reduced pressure, diluted in minimal amount of water (1 mL), and purified by reverse-phase HPLC to yield **2a** as a white solid (4.7 mg, 32%).

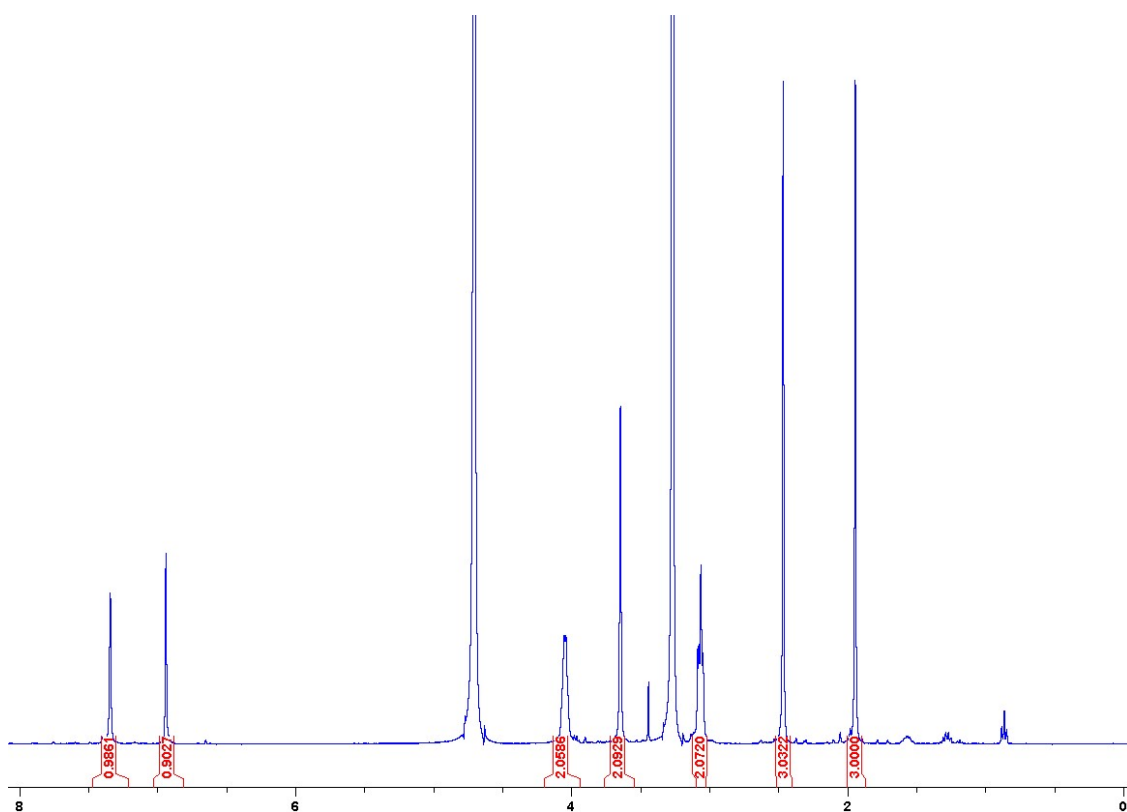
¹H NMR (400 MHz, D₂O): δ 7.34 (s, 1H, H-6''), 6.94 (s, 1H, H-5'), 4.03 (q, 2H, *J* = 6.8 Hz, H-1), 3.64 (s, 2H, H-7'), 3.06 (t, 2H, *J* = 6.8 Hz, H-2), 2.46 (s, 3H, H-7''), 1.94 (s, 3H, H-6').

¹³C NMR (100 MHz, D₂O) δ 165.7 (C-2''), 160.7 (C-4''), 140.5 (C-6''), 135.6 (C-4'), 134.8 (C-3'), 133.7 (C-2'), 120.7 (C-5'), 115.3 (C-5''), 66.6 (C-1), 28.9 (C-2), 27.0 (C-7'), 20.5 (C-7''), 11.3 (C-6').

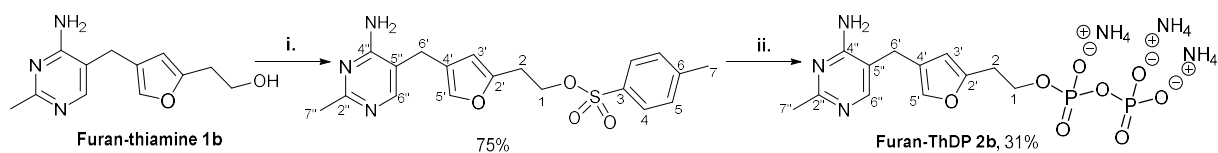
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₃H₁₉N₃O₇SP₂: 424.0497; found: 424.0499.

Analytical data are consistent with those previously reported.^{1,3}

¹H NMR in D₂O:



Preparation of furan-ThDP **2b**:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]furan-2-yl}ethyl 4-methylbenzene-1-sulfonate

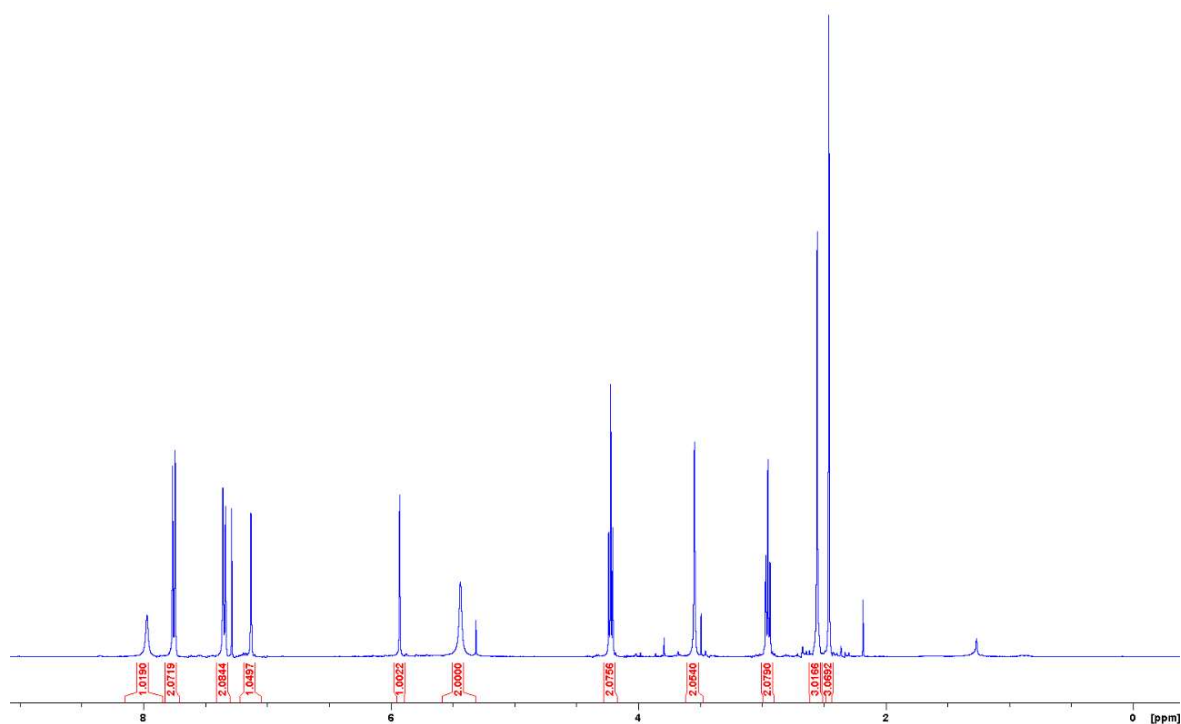
To a stirred solution of **1b** (47 mg, 0.2 mmol) in dry pyridine (1 mL, 0.2 M) under nitrogen at 0 °C was added *p*-TsCl (190 mg, 1 mmol). The resultant mixture was stirred at 0 °C for 4 h, quenched with cold 1 M HCl (1 mL), diluted with water (15 mL), neutralised with sodium bicarbonate to pH 7, and extracted with DCM (70 mL). The organic phase was washed with sat. aq. Cu₂SO₄, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield **1b**'s tosylate as a brown solid (58 mg, 75%).

¹H NMR (400 MHz, CDCl₃): δ 7.97 (s, 1H, H-6''), 7.76 (d, 2H, *J* = 8.0 Hz, H-4), 7.35 (d, 2H, *J* = 8.0 Hz, H-5), 7.13 (s, 1H, H-5'), 5.93 (s, 1H, H-3'), 5.44 (br, NH₂), 4.22 (t, 2H, *J* = 6.7 Hz, H-1), 3.55 (s, 2H, H-6'), 2.95 (t, 2H, *J* = 6.7 Hz, H-2), 2.56 (s, 3H, H-7''), 2.46 (s, 3H, H-7).

¹³C NMR (100 MHz, CDCl₃) δ 165.1 (C-2''), 162.0 (C-4''), 152.5 (C-6''), 151.7 (C-2'), 144.9 (C-3), 139.1 (C-5'), 132.8 (C-6), 129.9 (C-5), 128.1 (C-4), 121.1 (C-4'), 112.1 (C-5''), 108.5 (C-3'), 67.8 (C-1), 28.4 (C-2), 24.6 (C-7''), 24.2 (C-6'), 21.6 (C-7).

Analytical data are consistent with those previously reported.⁷

¹H NMR of **1b**'s tosylate in CDCl₃:



Furan-thiamine diphosphate (Furan-ThDP) 2b

To a stirred solution of **1b**'s tosylate (39 mg, 0.1 mmol) in dry MeCN (0.5 mL, 0.2 M) under nitrogen at 0 °C was added tris(tetra-*n*-butylammonium) hydrogen diphosphate (180 mg, 0.2 mmol). The resultant mixture was stirred at 0 °C overnight, concentrated under reduced pressure, diluted in minimal amount of water (1 mL), and purified by reverse-phase HPLC to yield **2b** as a white solid (14 mg, 32%).

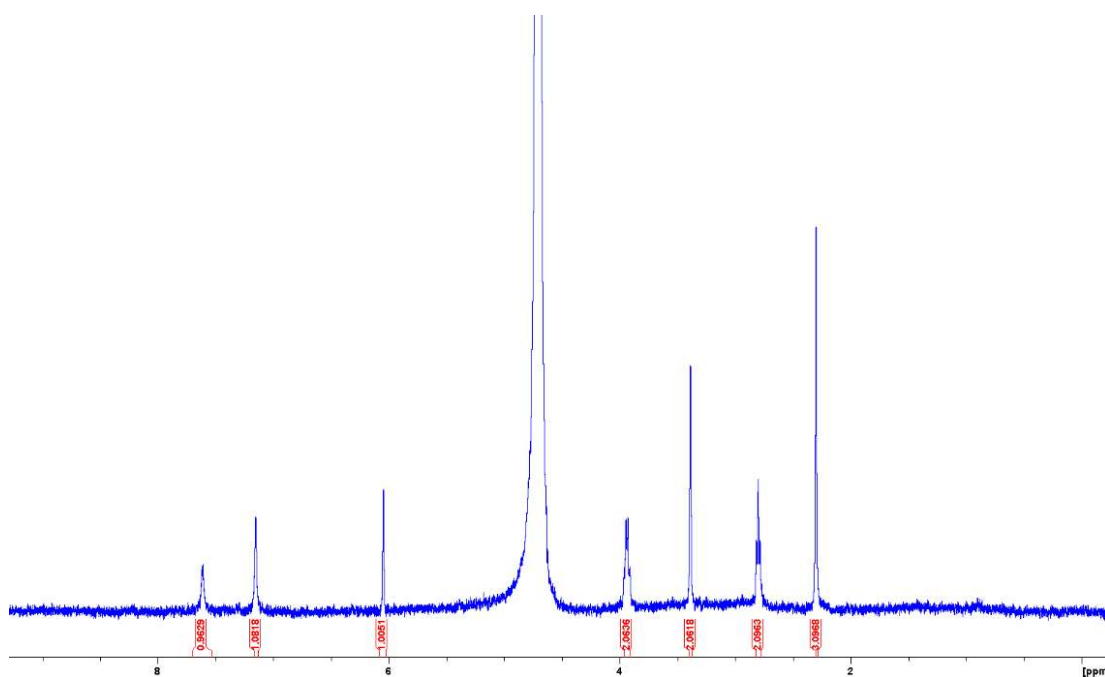
¹H NMR (400 MHz, D₂O): δ 7.62 (s, 1H, H-6''), 7.13 (s, 1H, H-5'), 6.04 (s, 1H, H-3'), 3.95 (q, 2H, *J* = 8.0 Hz, H-1), 3.41 (s, 2H, H-6'), 2.78 (t, 2H, *J* = 8.0 Hz, H-2), 2.30 (s, 3H, H-7'').

¹³C NMR (100 MHz, D₂O) δ 164.0 (C-2''), 161.2 (C-4''), 155.5 (C-6''), 154.0 (C-2'), 139.1 (C-5'), 118.7 (C-4'), 114.7 (C-5''), 107.7 (C-3'), 63.6 (C-1), 28.9 (C-2), 22.7 (C-6'), 21.5 (C-7'').

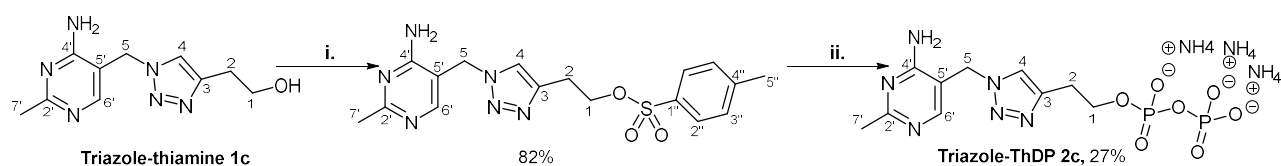
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₂H₁₇N₃O₈P₂: 394.0568; found: 394.0577.

Analytical data are consistent with those previously reported.⁷

¹H NMR in D₂O:



Preparation of triazole-ThDP 2c:



2-[1-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-1H-1,2,3-triazol-4-yl]ethyl 4-methylbenzene-1-sulfonate

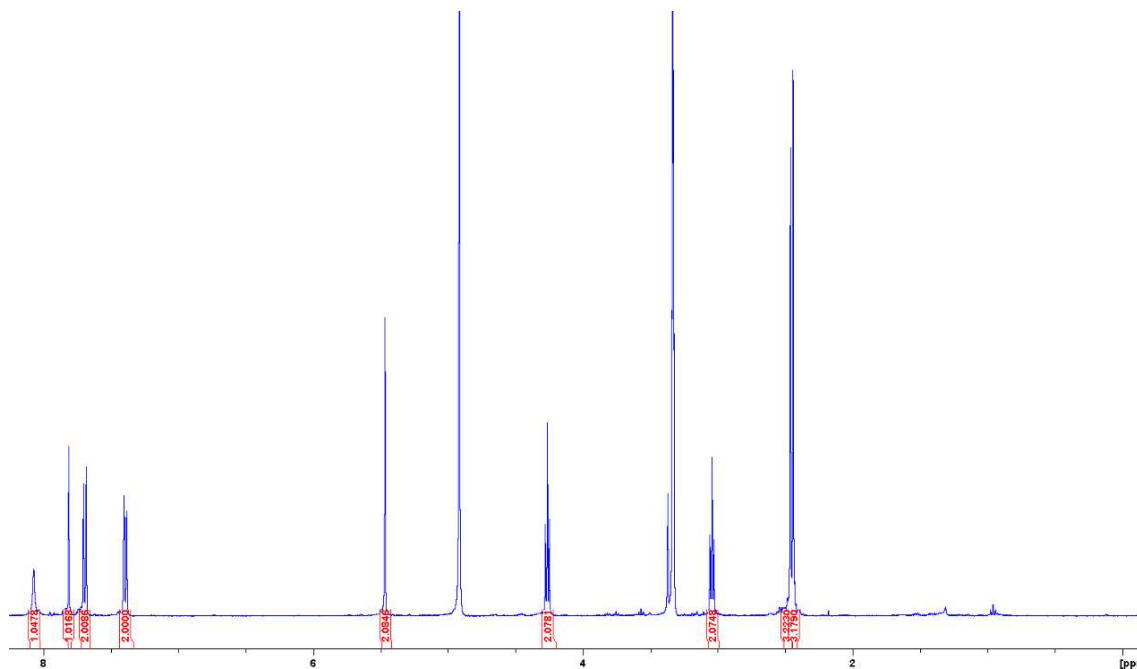
To a stirred solution of **1c** (234 mg, 1 mmol) in dry pyridine (5 mL, 0.2 M) under nitrogen at 0 °C was added *p*-TsCl (953 mg, 5 mmol). The resultant mixture was stirred at r.t. for 4 h, quenched with cold 1 M HCl (5 mL), diluted with water (15 mL), neutralised with sodium bicarbonate to pH 7, and extracted with DCM (100 mL). The organic phase was washed with sat. aq. Cu₂SO₄, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% MeOH in DCM) to yield **1c**'s tosylate as a white solid (318 mg, 82%).

¹H NMR (400 MHz, CD₃OD): δ 8.07 (s, 1H, H-6'), 7.79 (s, 1H, H-4), 7.67 (d, 2H, $J = 8.1$ Hz, H-2''), 7.36 (d, 2H, $J = 8.1$ Hz, H-3''), 5.45 (s, 2H, H-5), 4.25 (t, 2H, $J = 6.3$ Hz, H-1), 3.03 (t, 2H, $J = 6.3$ Hz, H-2), 2.42 (s, 3H, H-7'), 2.41 (s, 3H, H-5'').

¹³C NMR (100 MHz, CD₃SOCD₃) δ 166.8 (C-2'), 161.8 (C-4'), 156.5 (C-6'), 145.1 (C-3), 142.3 (C-1''), 133.3 (C-4''), 130.4 (C-3''), 127.7 (C-2''), 123.2 (C-4), 108.7 (C-5'), 69.5 (C-1), 46.3 (C-5), 25.8 (C-2), 25.5 (C-7'), 21.0 (C-5'').

Analytical data are consistent with those previously reported.⁹

¹H NMR of **1c**'s tosylate in CD₃OD:



Triazole-thiamine diphosphate (Triazole-ThDP) 2c

To a stirred solution of **1c**'s tosylate (39 mg, 0.1 mmol) in dry MeCN (0.5 mL, 0.2 M) under nitrogen at 0 °C was added tris(tetra-*n*-butylammonium) hydrogen diphosphate (186 mg, 0.2 mmol). The resultant mixture was stirred at r.t. overnight, concentrated under reduced pressure, diluted in minimal amount of water (1 mL), and purified by reverse-phase HPLC to yield **2c** as a white solid (12 mg, 27%).

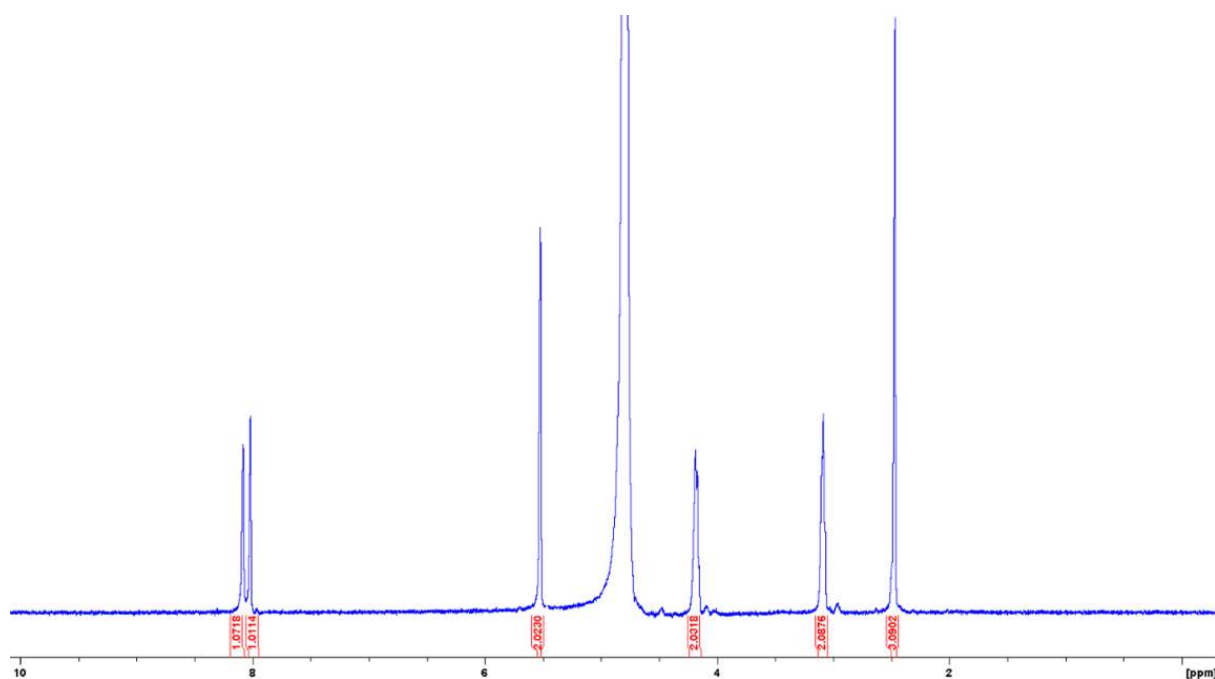
¹H NMR (400 MHz, D₂O): δ 8.08 (s, 1H, H-6'), 8.02 (s, 1H, H-4), 5.53 (s, 2H, H-5), 4.18 (m, 2H, H-1), 3.09 (t, 2H, *J* = 5.6 Hz, H-2), 2.47 (s, 3H, H-7').

¹³C NMR (100 MHz, D₂O) δ 165.6 (C-6'), 163.2 (C-2'), 162.5 (C-4'), 143.2 (C-3), 125.2 (C-4), 109.3 (C-5'), 64.6 (C-1), 46.9 (C-5), 26.2 (C-2), 21.0 (C-7').

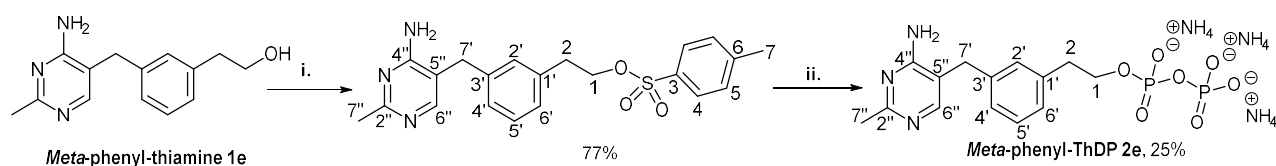
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₀H₁₆N₆O₇P₂: 395.0633; found: 395.0622.

Analytical data are consistent with those previously reported.⁹

¹H NMR in D₂O:



Preparation of *meta*-phenyl-ThDP **2e**:



2-[3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl]ethyl 4-methylbenzene-1-sulfonate

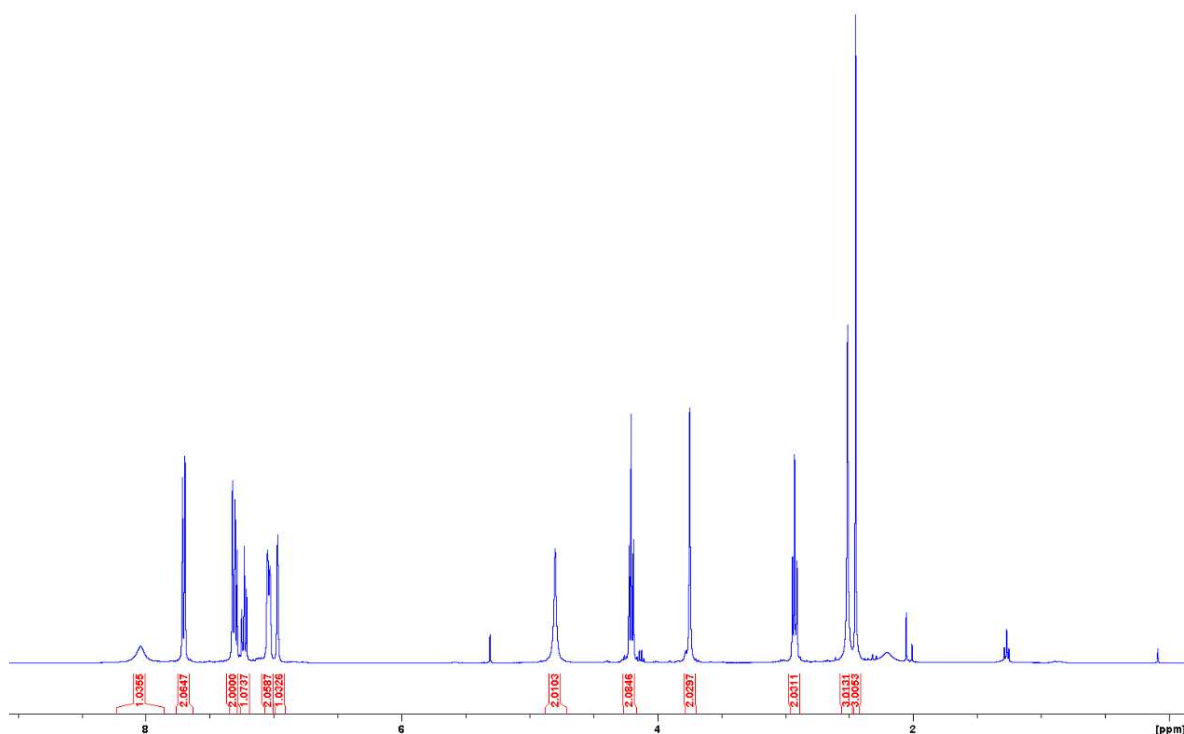
To a stirred solution of **1e** (73 mg, 0.3 mmol) in dry pyridine (1.5 mL, 0.2 M) under nitrogen at 0 °C was added *p*-TsCl (286 mg, 1.5 mmol). The resultant mixture was stirred at r.t. for 4 h, quenched with cold 1 M HCl (2 mL), diluted with water (20 mL), neutralised with sodium bicarbonate to pH 7, and extracted with DCM (70 mL). The organic phase was washed with sat. aq. Cu₂SO₄, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% MeOH in DCM) to yield **1e**'s tosylate as a brown solid (92 mg, 77%).

¹H NMR (400 MHz, CDCl₃): δ 8.02 (s, 1H, H-6''), 7.70 (d, 2H, *J* = 8.2 Hz, H-4), 7.31 (d, 2H, *J* = 8.2 Hz, H-5), 7.23 (dd, 1H, *J* = 7.5 and 7.8 Hz, H-5'), 7.03-7.04 (m, 2H, H-2' and H-4'), 6.97 (m, 1H, H-6'), 4.80 (br, NH₂), 4.20 (t, 2H, *J* = 6.8 Hz, H-1), 3.75 (s, 2H, H-7'), 2.92 (t, 2H, *J* = 6.8 Hz, H-2), 2.52 (s, 3H, H-7''), 2.45 (s, 3H, H-7).

¹³C NMR (100 MHz, CDCl₃) δ 166.6 (C-2''), 161.5 (C-4''), 155.5 (C-6''), 144.7 (C-3), 137.6 (C-3'), 137.2 (C-1'), 133.0 (C-6), 129.8 (C-5), 129.3 (C-5'), 128.9 (C-6'), 127.8 (C-4), 127.6 (C-2'), 126.8 (C-4'), 111.6 (C-5''), 70.4 (C-1), 35.4 (C-2), 34.5 (C-7'), 25.6 (C-7''), 21.6 (C-7).

Analytical data are consistent with those previously reported.²

¹H NMR of **1e**'s tosylate in CDCl₃:



Meta-phenyl-thiamine diphosphate (meta-phenyl-ThDP) 2e

To a stirred solution of **1e**'s tosylate (40 mg, 0.1 mmol) in dry MeCN (0.5 mL, 0.2 M) under nitrogen at 0 °C was added tris(tetra-*n*-butylammonium) hydrogen diphosphate (186 mg, 0.2 mmol). The resultant mixture was stirred at r.t. overnight, concentrated under reduced pressure, diluted in minimal amount of water (1 mL), and purified by reverse-phase HPLC to yield **2e** as a white solid (12 mg, 26%).

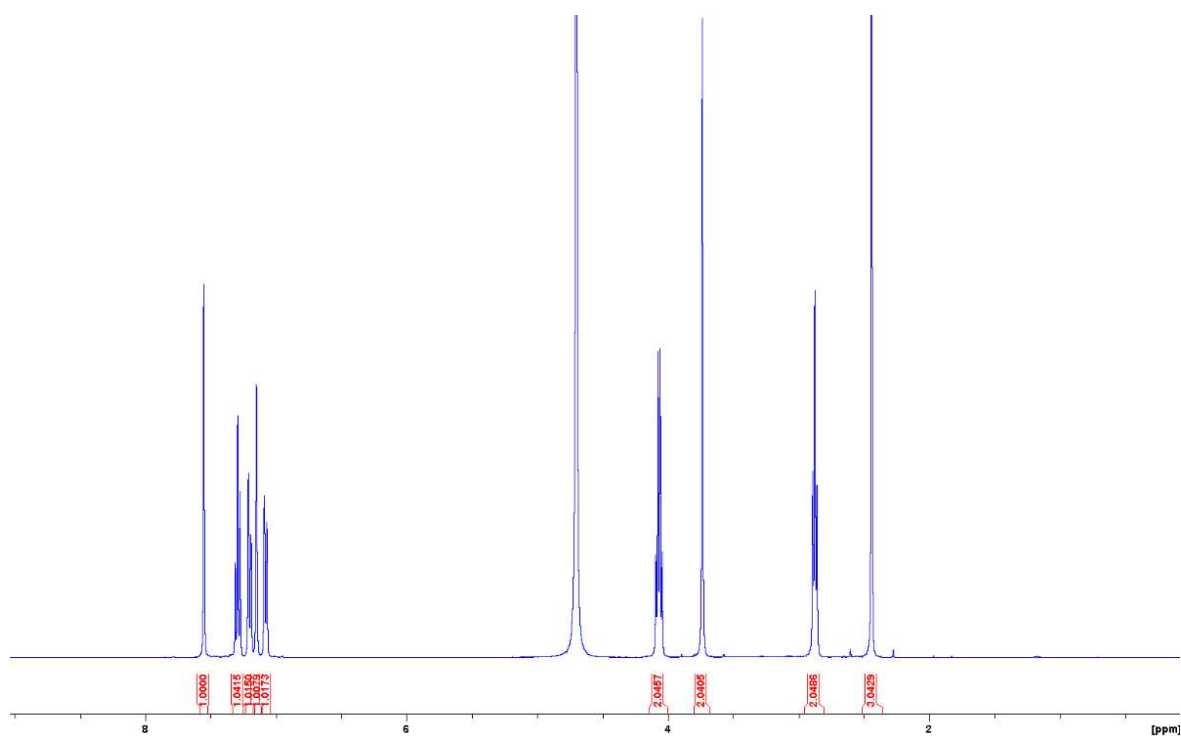
¹H NMR (400 MHz, D₂O): δ 7.56 (s, 1H, H-6''), 7.30 (t, 1H, *J* = 7.5 Hz, H-5'), 7.22 (d, 1H, *J* = 7.5 Hz, H-4'), 7.18 (s, 1H, H-2'), 7.08 (d, 1H, *J* = 7.5 Hz, H-6'), 4.07 (q, 2H, *J* = 6.6 Hz, H-1), 3.75 (s, 2H, H-7'), 2.88 (t, 2H, *J* = 6.6 Hz, H-2), 2.45 (s, 3H, H-7'').

¹³C NMR (100 MHz, D₂O) δ 164.0 (C-2''), 160.7 (C-4''), 141.1 (C-6''), 139.3 (C-3'), 135.3 (C-1'), 129.8 (C-4'), 129.2 (C-6'), 128.0 (C-2'), 126.8 (C-5'), 115.6 (C-5''), 67.0 (C-1), 35.9 (C-2), 32.5 (C-7'), 20.4 (C-7'').

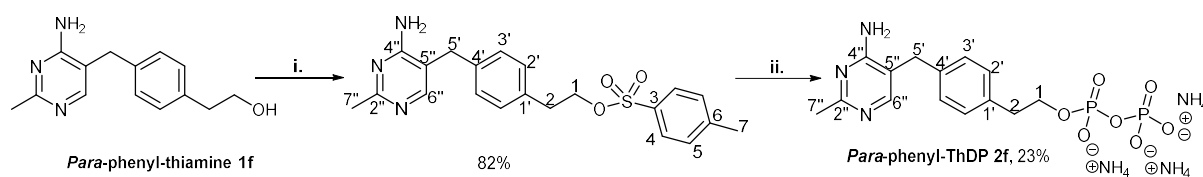
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₄H₁₉N₃O₇P₂: 404.0776; found: 404.0795.

Analytical data are consistent with those previously reported.²

¹H NMR in D₂O:



Preparation of *meta*-phenyl-ThDP **2f**:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl}ethyl 4-methylbenzene-1-sulfonate

To a stirred solution of **1f** (73 mg, 0.3 mmol) in dry pyridine (1.5 mL, 0.2 M) under nitrogen at 0 °C was added *p*-TsCl (286 mg, 1.5 mmol). The resultant mixture was stirred at r.t. for 4 h, quenched with cold 1 M HCl (2 mL), diluted with water (20 mL), neutralised with sodium bicarbonate to pH 7, and extracted with DCM (70 mL). The organic phase was washed with sat. aq. Cu₂SO₄, dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (10% MeOH in DCM) to yield **1f**'s tosylate as a brown solid (97 mg, 82%).

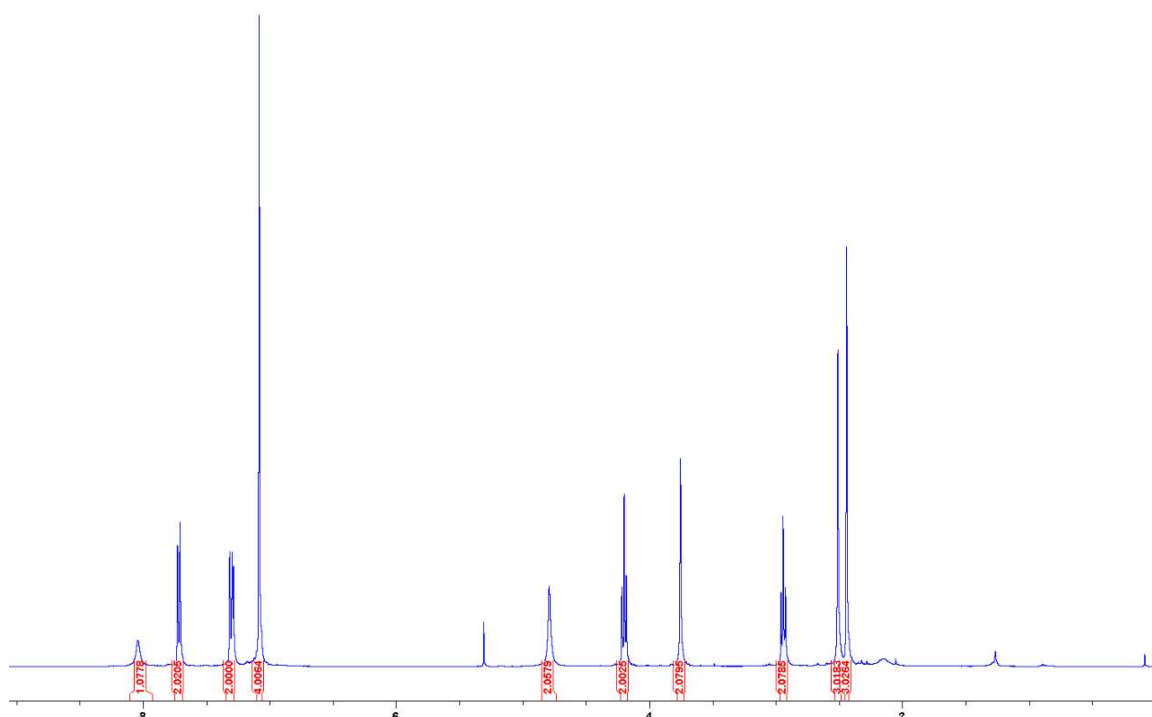
m.p. 118-120 °C.

¹H NMR (400 MHz, CDCl₃): δ 8.04 (s, 1H, H-6''), 7.72 (d, 2H, *J* = 8.3 Hz, H-4), 7.31 (d, 2H, *J* = 8.3 Hz, H-5), 7.09 (m, 4H, H-2' and H-3'), 4.79 (br, NH₂), 4.20 (t, 2H, *J* = 6.9 Hz, H-1), 3.75 (s, 2H, H-5'), 2.94 (t, 2H, *J* = 6.9 Hz, H-2), 2.52 (s, 3H, H-7''), 2.44 (s, 3H, H-7).

¹³C NMR (100 MHz, CDCl₃) δ 166.4 (C-2''), 161.6 (C-4''), 155.9 (C-6''), 144.8 (C-3), 135.8 (C-4'), 135.1 (C-1'), 132.9 (C-6), 129.9 (C-5), 129.6 (C-3'), 128.5 (C-2'), 127.8 (C-4), 111.5 (C-5''), 70.6 (C-1), 34.9 (C-2), 34.8 (C-5'), 25.5 (C-7''), 21.7 (C-7).

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₁H₂₃N₃O₃S: 398.1538; found: 398.1515.

¹H NMR of **1f**'s tosylate in CDCl₃:



Para-phenyl-thiamine diphosphate (para-phenyl-ThDP) 2f

To a stirred solution of **1f**'s tosylate (40 mg, 0.1 mmol) in dry MeCN (0.5 mL, 0.2 M) under nitrogen at 0 °C was added tris(tetra-*n*-butylammonium) hydrogen diphosphate (186 mg, 0.2 mmol). The resultant mixture was stirred at r.t. overnight, concentrated under reduced pressure, diluted in minimal amount of water (1 mL), and purified by reverse-phase HPLC to yield **2f** as a white solid (11 mg, 23%).

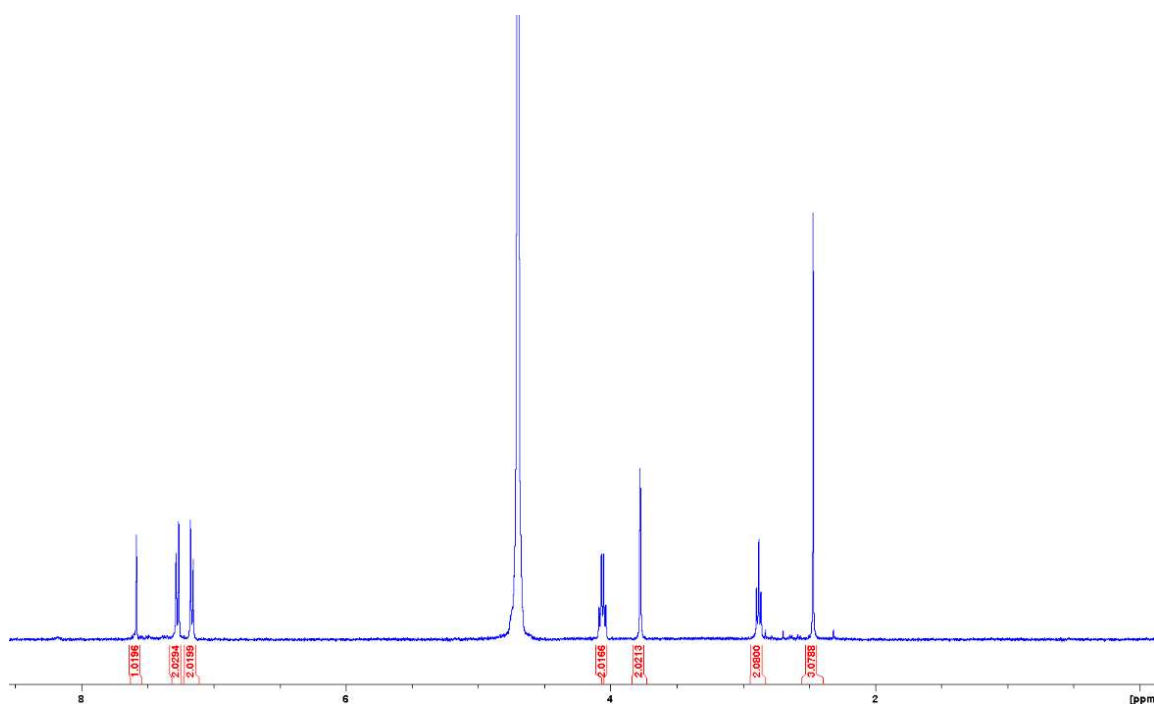
¹H NMR (400 MHz, D₂O): δ 7.58 (s, 1H, H-6''), 7.29 (d, 2H, *J* = 8.3 Hz, H-3'), 7.16 (d, 2H, *J* = 8.3 Hz, H-2'), 4.06 (q, 2H, *J* = 6.8 Hz, H-1), 3.75 (s, 2H, H-5'), 2.90 (t, 2H, *J* = 6.8 Hz, H-2), 2.49 (s, 3H, H-7'').

¹³C NMR (100 MHz, D₂O) δ 163.1 (C-2''), 160.8 (C-4''), 140.8 (C-6''), 137.3 (C-4'), 135.0 (C-1'), 129.8 (C-3'), 128.8 (C-2'), 115.8 (C-5''), 68.1 (C-1), 38.3 (C-2), 32.6 (C-5'), 20.5 (C-7'').

³¹P NMR (162 MHz, D₂O) δ -0.3 (d, *J* = 22 Hz), -9.0 (d, *J* = 22 Hz).

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₄H₁₉N₃O₇P₂: 404.0776; found: 404.0795.

¹H NMR in D₂O:

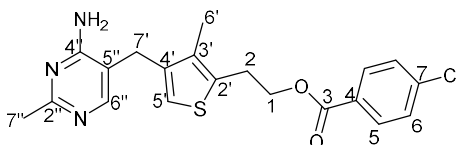


Preparation of *p*-chlorobenzoyl thiamine-ester analogues **3a-f**:

General procedure:

To a stirred solution of 4-chlorobenzoic acid (203 mg, 1.3 mmol) and DCC (618 mg, 3 mmol) in dry DMF (10 mL, 0.1 M) under nitrogen at 0 °C was added DMAP (158 mg, 1.3 mmol) and alcohol **1a-f** (1 mmol). The resultant mixture was stirred at r.t. for 2 days, diluted with DCM, filtered (to remove DCC/DCU), treated with aqueous phosphate buffer (pH 7) (50 mL), and extracted with DCM (200 mL). The organic phase was dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield ester **3a-f** as a solid.

2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl}ethyl 4-chlorobenzoate **3a**



Prepared from **1a**. Brown solid (200 mg, 50%).

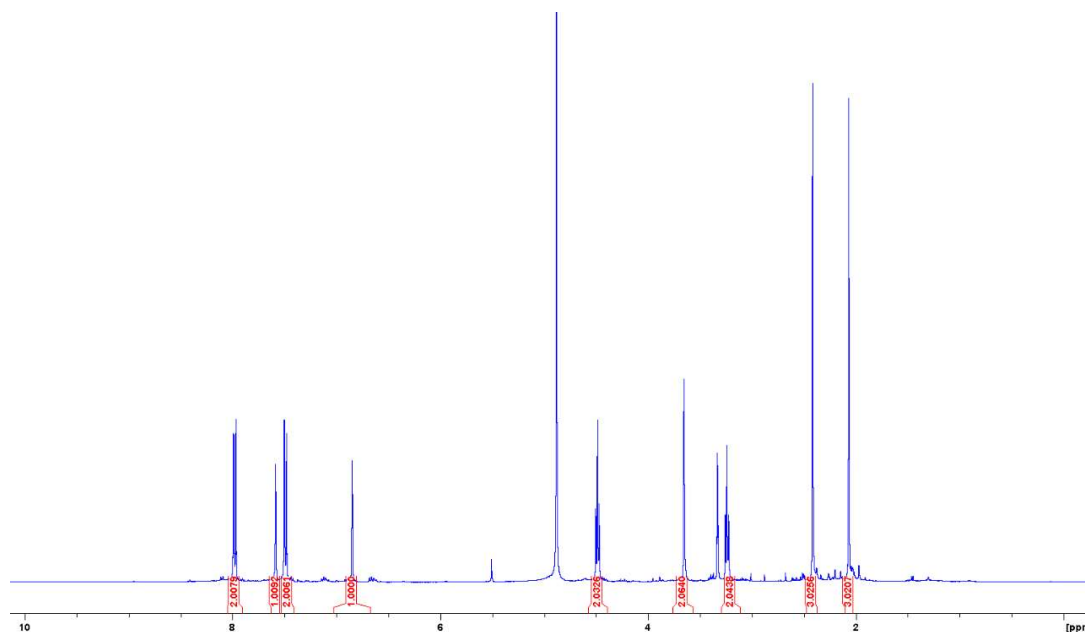
m.p. 173-174 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.98 (d, 2H, *J* = 8.6 Hz, H-5), 7.58 (s, 1H, H-6''), 7.49 (d, 2H, *J* = 8.6 Hz, H-6), 6.84 (s, 1H, H-5'), 4.48 (d, 2H, *J* = 6.5 Hz, H-1), 3.64 (s, 2H, H-7'), 3.24 (t, 2H, *J* = 6.5 Hz, H-2), 2.41 (s, 3H, H-7''), 2.07 (s, 3H, H-6').

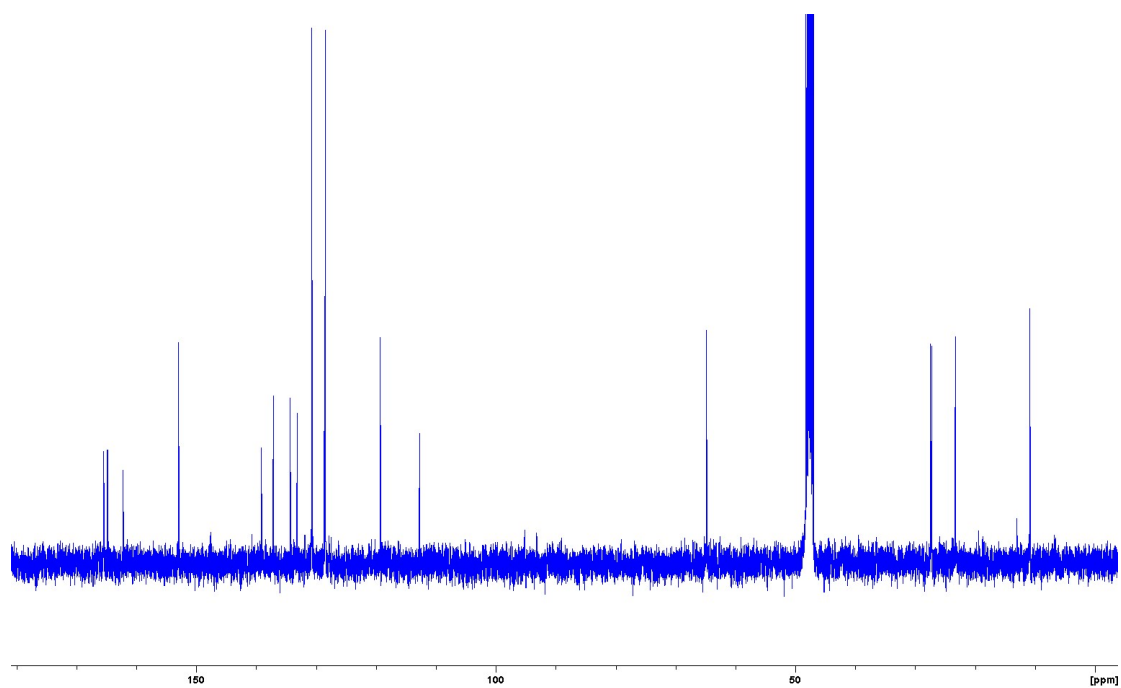
¹³C NMR (100 MHz, CD₃OD) δ 165.5 (C-3), 164.8 (C-2''), 162.2 (C-4''), 153.1 (C-5'), 139.2 (C-7), 137.2 (C-4'), 134.3 (C-3'), 133.2 (C-2'), 130.8 (C-5), 128.7 (C-4), 128.5 (C-6), 119.4 (C-5'), 112.5 (C-5''), 65.0 (C-1), 27.5 (C-7'), 27.4 (C-2), 23.3 (C-7''), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₀ClN₃O₂S: 402.1042; found: 402.1050.

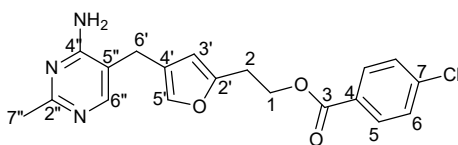
¹H NMR in CD₃OD:



^{13}C NMR in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]furan-2-yl}ethyl 4-chlorobenzoate **3b**



Prepared from **1b**. Brown solid (223 mg, 60%).

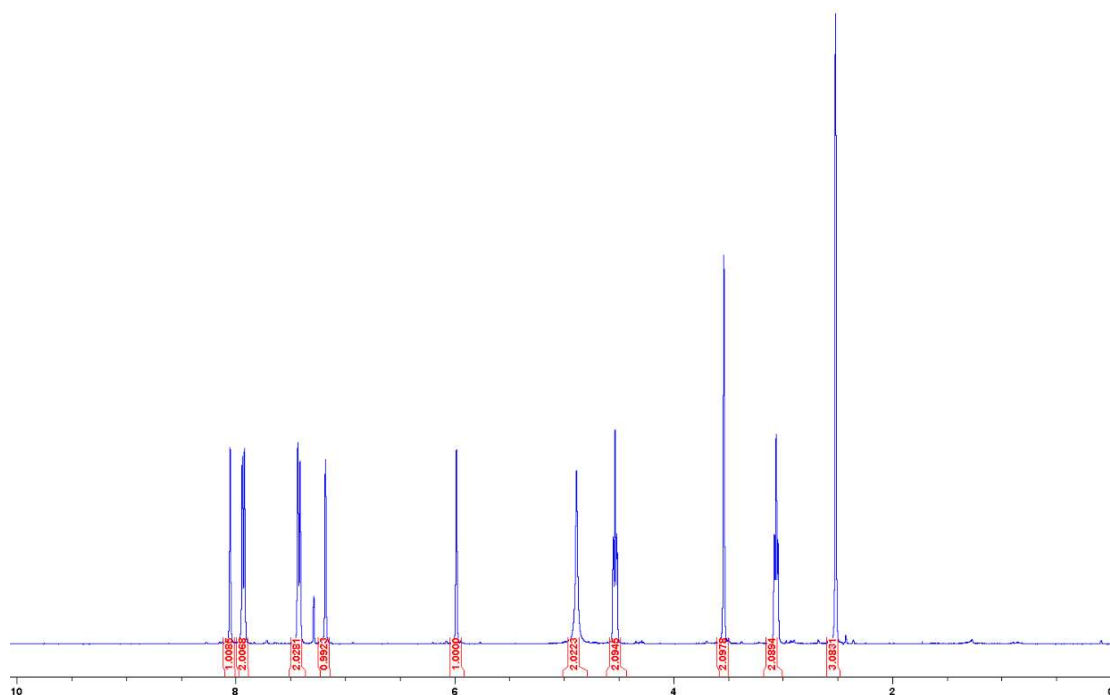
¹H NMR (400 MHz, CDCl₃) δ 8.06 (s, 1H, H-6''), 7.93 (d, 2H, *J* = 7.8 Hz, H-5), 7.42 (d, 2H, *J* = 7.8 Hz, H-6), 7.18 (s, 1H, H-5'), 6.00 (s, 1H, H-3'), 4.88 (br, NH₂), 4.53 (d, 2H, *J* = 6.5 Hz, H-1), 3.54 (s, 2H, H-6'), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.52 (s, 3H, H-7'').

¹³C NMR (100 MHz, CDCl₃) δ 166.2 (C-3), 165.3 (C-2''), 161.5 (C-4''), 155.5 (C-6''), 153.6 (C-2'), 139.3 (C-7), 138.5 (C-5'), 131.3 (C-5), 128.8 (C-6), 128.5 (C-4), 121.6 (C-4'), 112.2 (C-5''), 107.5 (C-3'), 63.0 (C-1), 28.0 (C-2), 25.2 (C-7''), 24.3 (C-6').

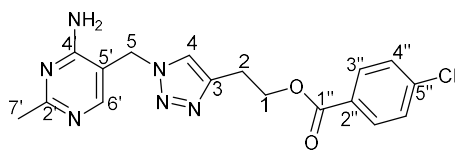
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₉H₁₈ClN₃O₃: 372.1114; found: 372.1122.

Analytical data are consistent with those previously reported.⁸

¹H NMR in CDCl₃:



2-[1-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-1H-1,2,3-triazol-4-yl]ethyl 4-chlorobenzoate **3c**



Prepared from **1c**. White solid (197 mg, 53%).

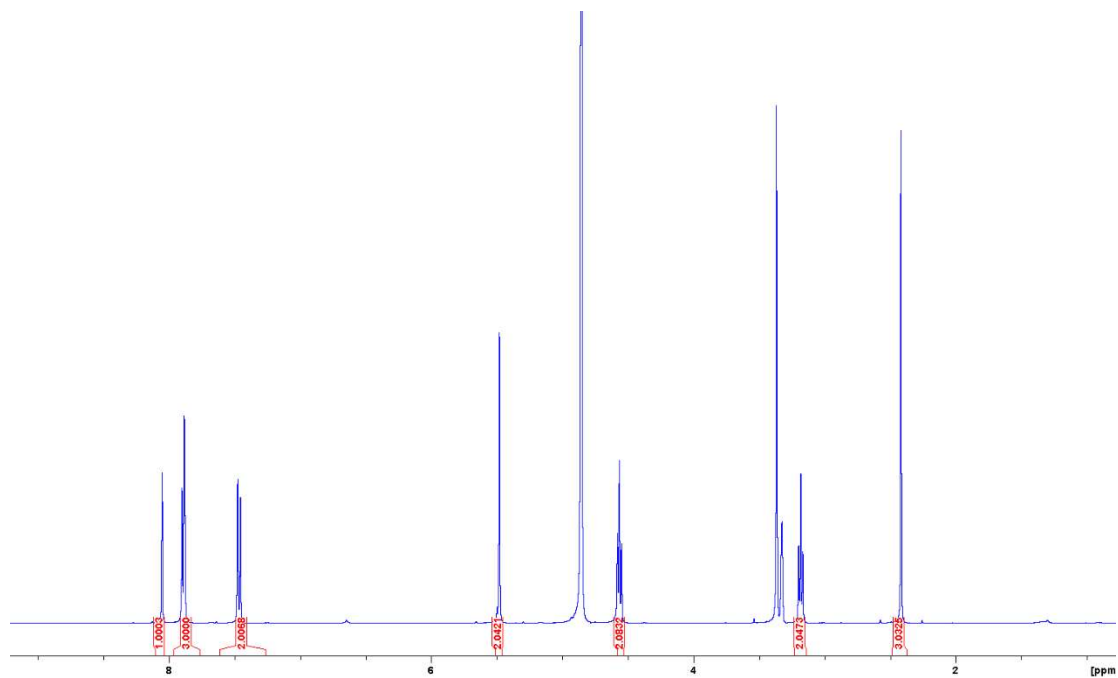
$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 8.05 (s, 1H, H-6'), 7.92 (m, 3H, H-3'' and H-4), 7.48 (d, 2H, $J = 8.4$ Hz, H-4''), 5.48 (s, 2H, H-5), 4.55 (t, 2H, $J = 6.6$ Hz, H-1), 3.18 (t, 2H, $J = 6.6$ Hz, H-2), 2.42 (s, 3H, H-7').

$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 167.7 (C-2'), 165.5 (C-1''), 162.0 (C-4'), 155.5 (C-6'), 144.5 (C-3), 139.0 (C-5''), 130.5 (C-3''), 128.6 (C-2''), 128.2 (C-4''), 122.2 (C-4), 108.8 (C-5'), 63.3 (C-1), 47.5 (C-5), 25.0 (C-2), 23.8 (C-7').

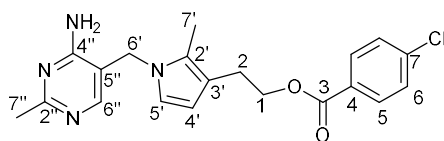
HRMS (ESI) m/z : $[\text{M}+\text{H}^+]$ calculated for $\text{C}_{17}\text{H}_{17}\text{ClN}_6\text{O}_2$: 373.1179; found: 373.1188.

Analytical data are consistent with those previously reported.⁸

$^1\text{H NMR}$ in CD_3OD :



2-{1-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-2-methyl-1H-pyrrol-3-yl}ethyl 4-chlorobenzoate **3d**



Prepared from **1d**. White solid (123 mg, 32%).

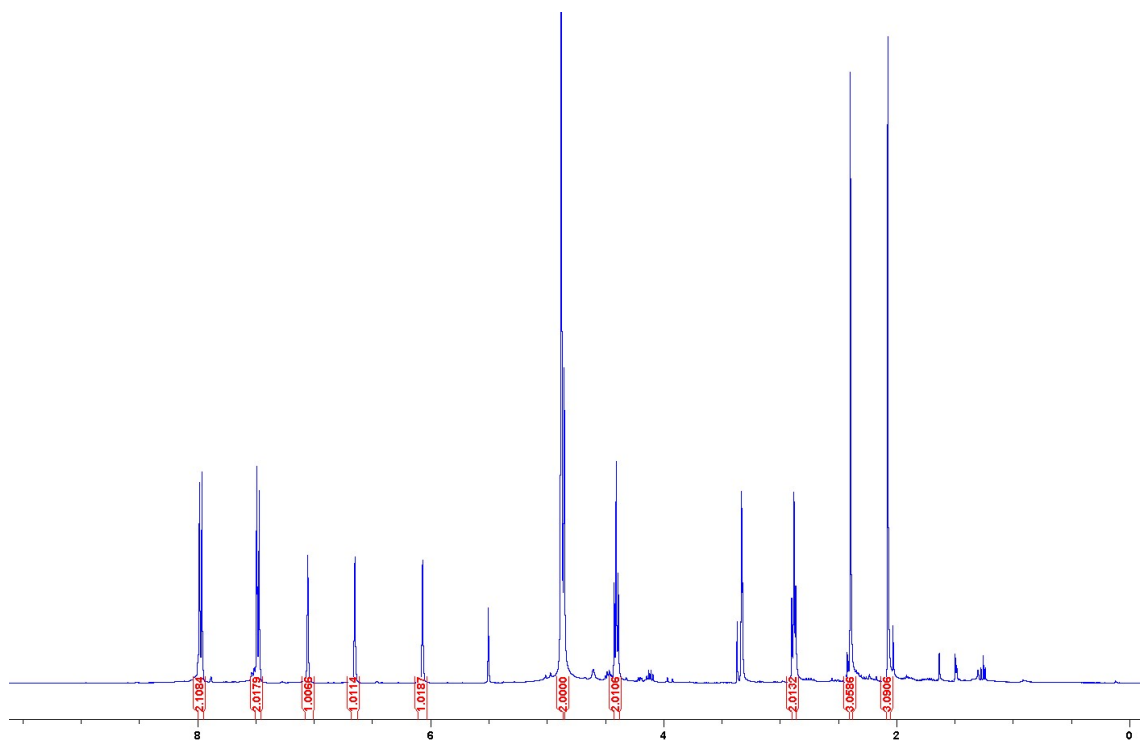
m.p. 180-181 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.97 (d, 2H, *J* = 8.7 Hz, H-5), 7.48 (d, 2H, *J* = 8.7 Hz, H-6), 7.06 (s, 1H, H-6''), 6.65 (d, 1H, *J* = 2.8 Hz, H-5'), 6.07 (d, 1H, *J* = 2.8 Hz, H-4'), 4.86 (s, 2H, H-6'), 4.41 (t, 2H, *J* = 7.1 Hz, H-1), 2.88 (t, 2H, *J* = 7.2 Hz, H-2), 2.40 (s, 3H, H-7''), 2.08 (s, 3H, H-7').

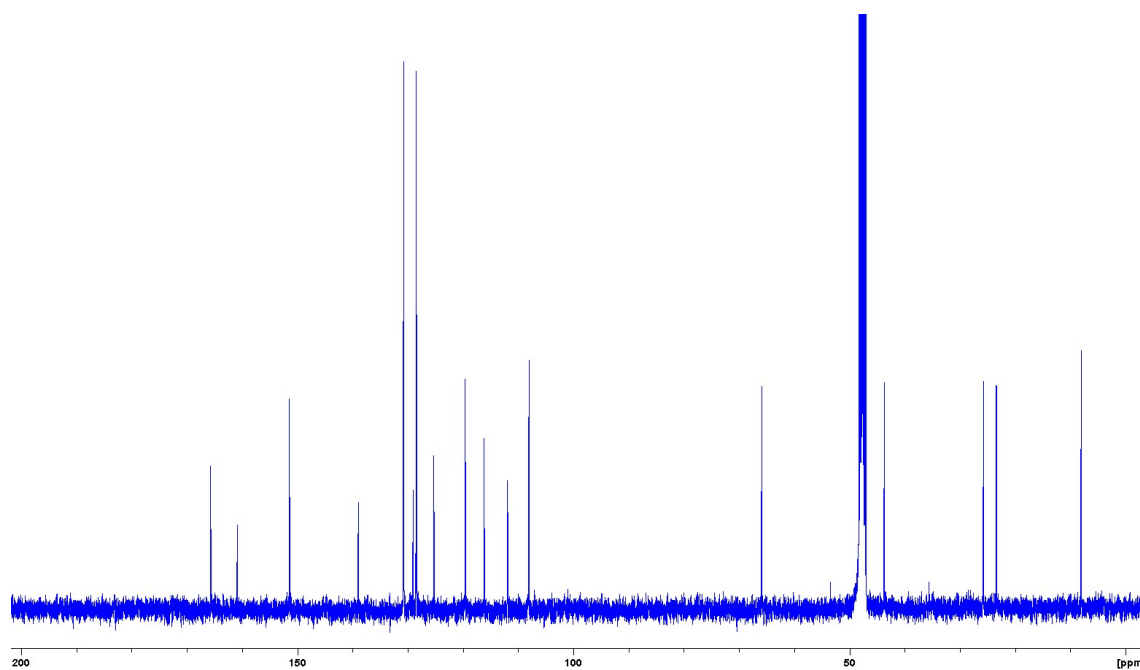
¹³C NMR (100 MHz, CD₃OD) δ 165.7 (C-2''), 165.7 (C-3), 160.9 (C-4''), 151.5 (C-6''), 139.0 (C-7), 130.5 (C-5), 128.9 (C-4), 128.4 (C-6), 125.3 (C-2'), 119.6 (C-5'), 116.2 (C-3'), 111.9 (C-5''), 108.1 (C-4'), 65.8 (C-1), 43.7 (C-6'), 25.7 (C-2), 23.4 (C-7''), 8.0 (C-7').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₁ClN₄O₂: 385.1431; found: 385.1437.

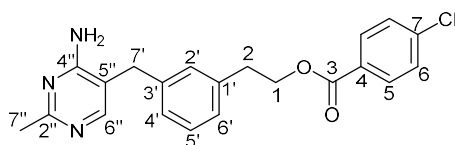
¹H NMR in CD₃OD:



^{13}C NMR in CD_3OD :



2-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl}ethyl 4-chlorobenzoate **3e**



Prepared from **1e**. Yellow solid (153 mg, 41%).

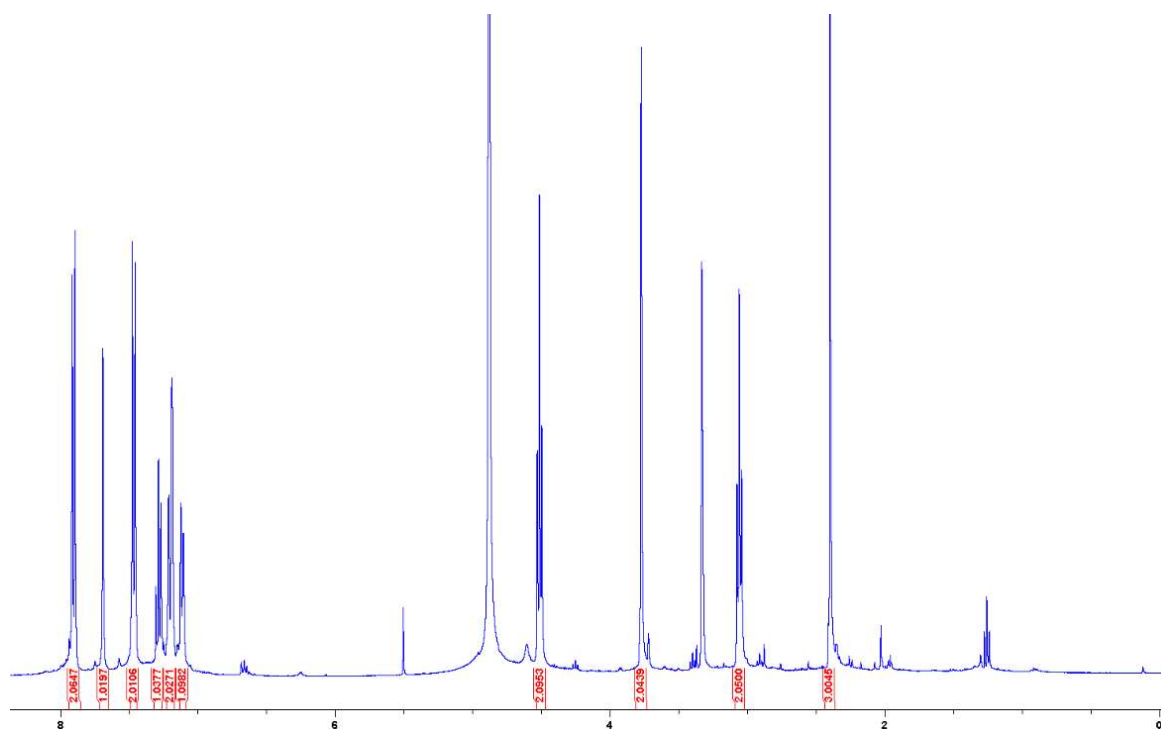
m.p. 185-187 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.69 (s, 1H, H-6''), 7.46 (d, 2H, *J* = 8.2 Hz, H-6), 7.28 (m, 1H, H-5'), 7.20 (m, 2H, H-2' and H-4'), 7.11 (d, 1H, *J* = 7.7 Hz, H-6'), 7.09 (d, 2H, *J* = 8.2 Hz, H-5), 4.51 (t, 2H, *J* = 6.7 Hz, H-1), 3.77 (s, 2H, H-7), 3.06 (t, 2H, *J* = 6.7 Hz, H-2), 2.40 (s, 3H, H-7'').

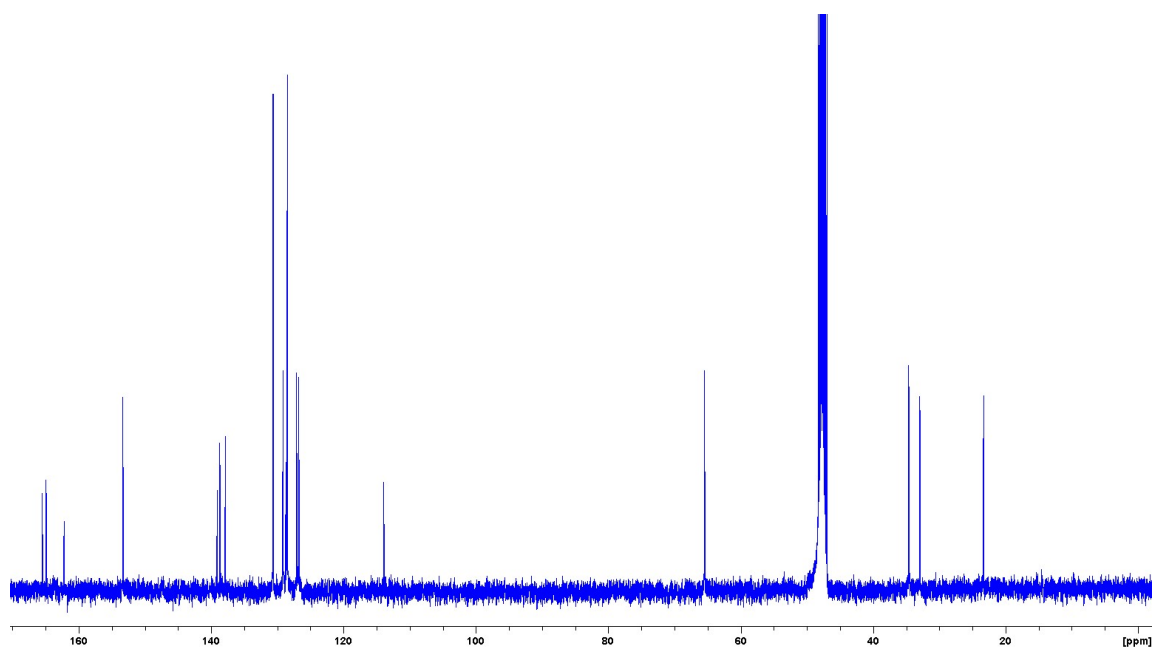
¹³C NMR (100 MHz, CD₃OD) δ 165.5 (C-3), 164.9 (C-2''), 162.2 (C-4''), 153.3 (C-6''), 139.0 (C-7), 138.7 (C-3'), 137.8 (C-1'), 130.6 (C-5), 129.1 (C-4'), 128.7 (C-4), 128.6 (C-6'), 128.5 (C-6), 127.0 (C-2'), 126.7 (C-5'), 113.9 (C-5''), 65.5 (C-1), 34.6 (C-2), 32.9 (C-7'), 23.3 (C7'').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₁H₂₀ClN₃O₂: 382.1322; found: 382.1333.

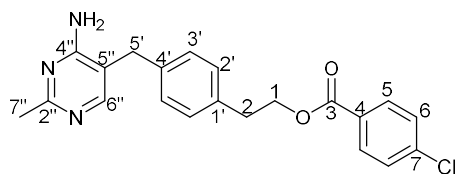
¹H NMR in CD₃OD:



^{13}C NMR in CD_3OD :



2-[4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]phenyl]ethyl 4-chlorobenzoate **3f**



Prepared from **1f**. White solid (183 mg, 48%).

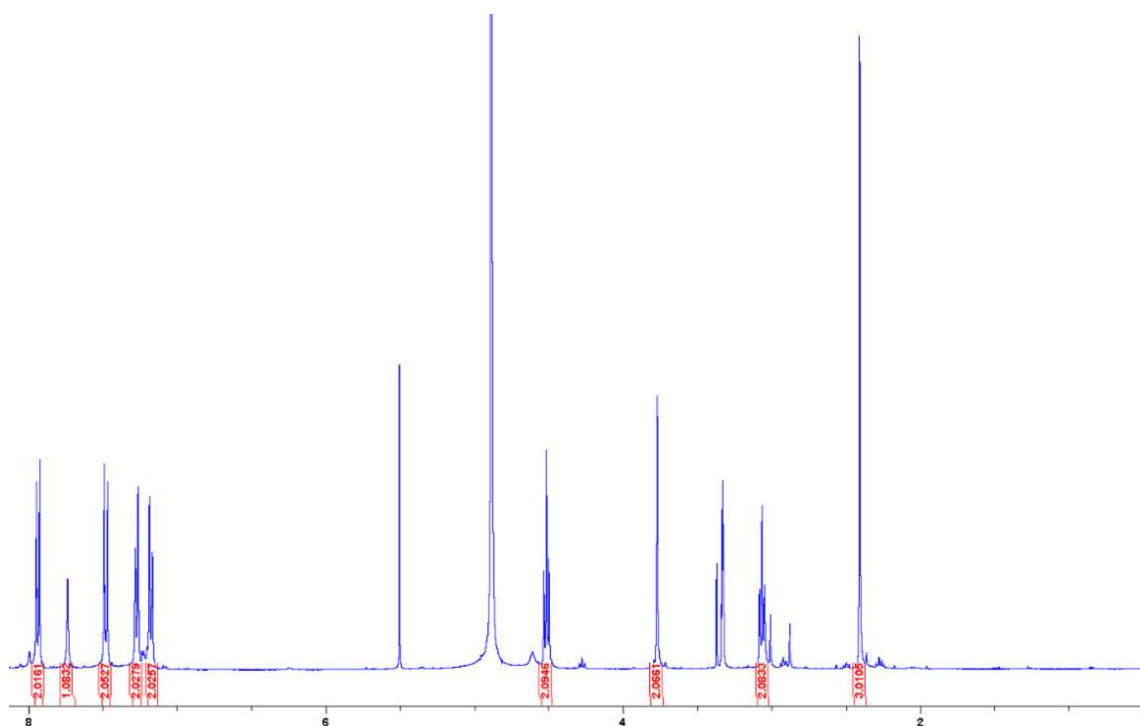
m.p. 199-201 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.94 (d, 2H, *J* = 8.8 Hz, H-5), 7.74 (s, 1H, H-6''), 7.48 (d, 2H, *J* = 8.8 Hz, H-6), 7.27 (d, 2H, *J* = 8.1 Hz, H-3'), 7.17 (d, 2H, *J* = 8.1 Hz, H-2'), 4.51 (t, 2H, *J* = 6.7 Hz, H-1), 3.77 (s, 2H, H-5'), 3.07 (t, 2H, *J* = 6.7 Hz, H-2), 2.41 (s, 3H, H-7'').

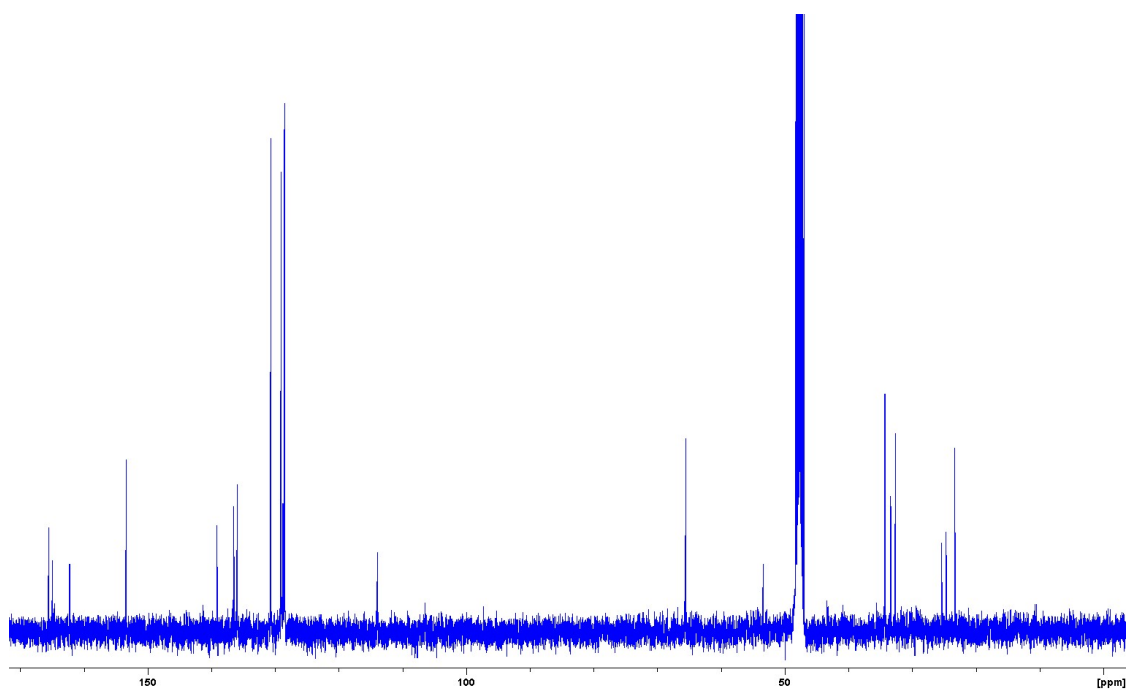
¹³C NMR (100 MHz, CD₃OD) δ 165.5 (C-3), 164.9 (C-2''), 162.2 (C-4''), 153.4 (C-6''), 139.1 (C-7), 136.5 (C-4'), 135.9 (C-1'), 130.6 (C-5), 129.0 (C-3'), 128.8 (C-4), 128.5 (C-2'), 128.4 (C-6), 113.9 (C-5''), 65.5 (C-1), 34.3 (C-2), 32.6 (C-5'), 23.3 (C-7'').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₁H₂₀ClN₃O₂: 382.1322; found: 382.1328.

¹H NMR in CD₃OD:



^{13}C NMR in CD_3OD :

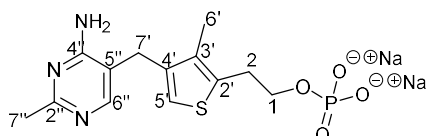


Preparation of thiamine monophosphate (ThMP) analogues 4a-f:

General procedure:

To a flask charged with corresponding **1a-f** (0.2 mmol) under nitrogen at 0 °C was added POCl₃ (1 mL, 0.2M) dropwise. The resultant mixture was stirred at 0 °C for 1 h and then at r.t. for 3 h. The reaction mixture was diluted with cold DCM (15 mL) at 0 °C, quenched with cold sat. aq. NaHCO₃ (10 mL) dropwise (NOTE: this quenching step is highly exothermic!), washed with DCM (2 x 20 mL), and concentrated under reduced pressure. The residue was diluted in minimal amount of water and purified by reverse-phase HPLC to yield **4a-f** as a solid.

Deazathiamine monophosphate (Deaza-ThMP) **4a**



Prepared from **1a**. White solid (19 mg, 27%).

m.p. 175-177 °C.

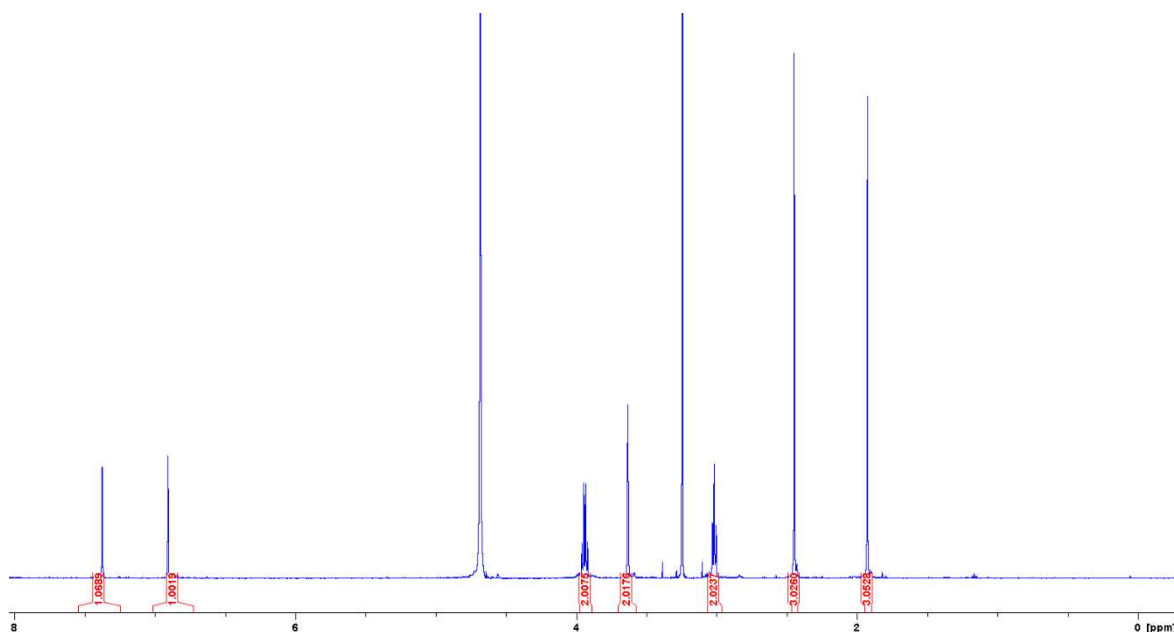
¹H NMR (400 MHz, D₂O) δ 7.37 (s, 1H, H-6''), 6.90 (s, 1H, H-5'), 3.93 (q, 2H, *J* = 6.5 Hz, H-1), 3.63 (s, 2H, H-7'), 3.01 (t, 2H, *J* = 6.5 Hz, H-2), 2.45 (s, 3H, H-7''), 1.92 (s, 3H, H-6').

¹³C NMR (100 MHz, D₂O) δ 164.0 (C-2''), 160.7 (C-4''), 140.5 (C-6''), 135.7 (C-4'), 134.9 (C-3'), 133.6 (C-2'), 120.6 (C-5'), 115.1 (C-5''), 65.6 (d, *J* = 5.4 Hz, C-1), 29.0 (C-2), 27.0 (C-7'), 20.5 (C-7''), 11.2 (C-6').

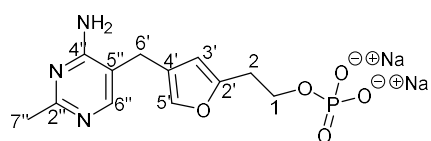
³¹P NMR (162 MHz, D₂O) δ 0.26.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₃H₁₈N₃O₄PS: 344.0833; found: 344.0835.

¹H NMR in D₂O:



Furan-thiamine monophosphate (Furan-ThMP) 4b



Prepared from **1b**. White solid (7 mg, 10%).

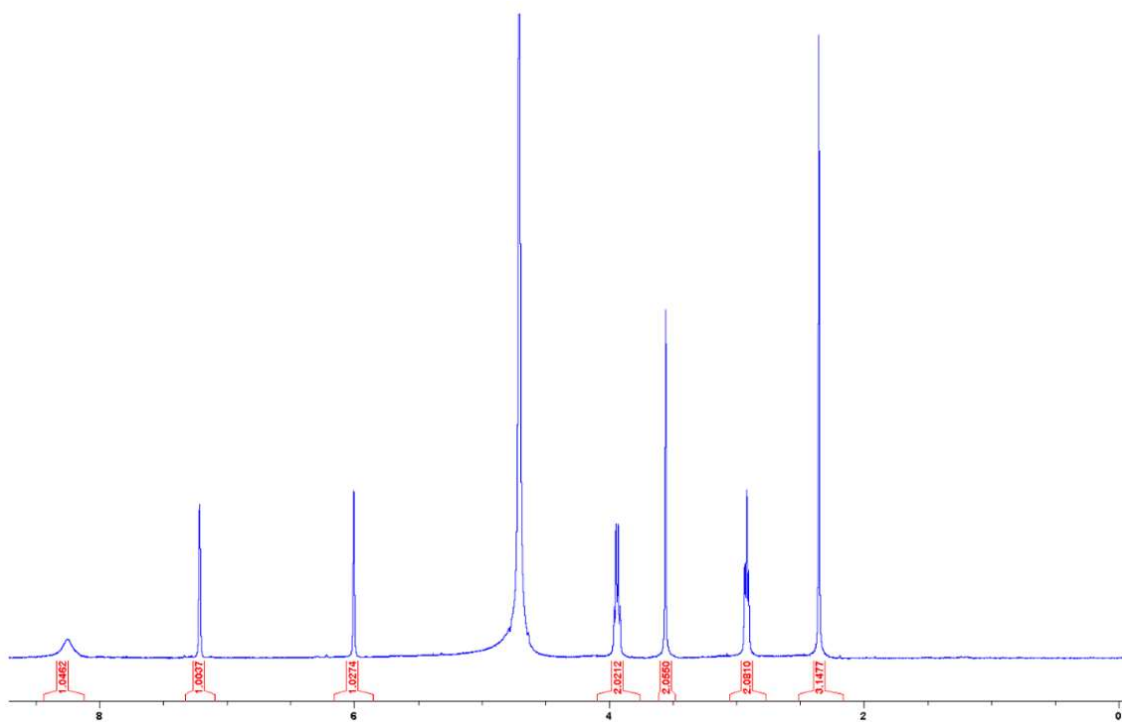
$^1\text{H NMR}$ (400 MHz, D_2O) δ 8.20 (s, 1H, H-6''), 7.21 (s, 1H, H-5'), 6.08 (s, 1H, H-3'), 3.93 (q, 2H, $J = 6.5$ Hz, H-1), 3.55 (s, 2H, H-6'), 2.88 (t, 2H, $J = 6.5$ Hz, H-2), 2.37 (s, 3H, H-7'').

$^{13}\text{C NMR}$ (100 MHz, D_2O) δ 163.8 (C-2''), 161.2 (C-4''), 157.1 (C-6''), 151.9 (C-2'), 139.2 (C-5'), 120.6 (C-4'), 113.6 (C-5''), 108.1 (C-3'), 63.5 (C-1), 29.5 (C-2), 25.2 (C-6'), 20.5 (C-7'').

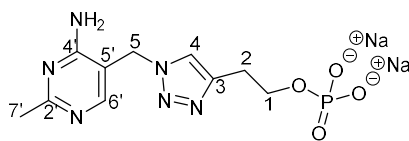
$^{31}\text{P NMR}$ (162 MHz, D_2O) δ 0.35.

HRMS (ESI) m/z : $[\text{M}+\text{H}^+]$ calculated for $\text{C}_{12}\text{H}_{16}\text{N}_3\text{O}_5\text{P}$: 314.0905; found: 314.0925.

$^1\text{H NMR}$ in D_2O :



Triazole-thiamine monophosphate (Triazole-ThMP) **4c**



Prepared from **1c**. White solid (14 mg, 20%).

m.p. 193-195 °C.

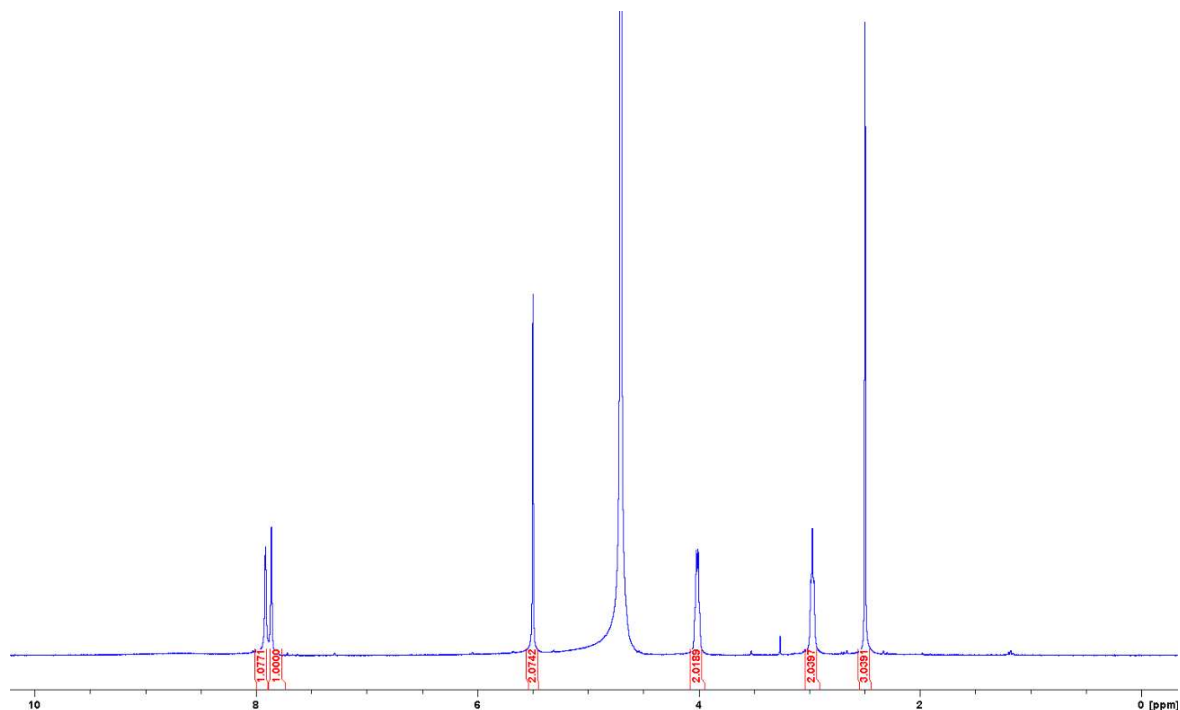
¹H NMR (400 MHz, D₂O): δ 7.92 (s, 1H, H-6'), 7.86 (s, 1H, H-4), 5.50 (s, 2H, H-5), 4.00 (q, 2H, *J* = 5.6 Hz, H-1), 2.98 (t, 2H, *J* = 5.6 Hz, H-2), 2.50 (s, 3H, H-7').

¹³C NMR (100 MHz, D₂O) δ 163.3 (C-2'), 162.8 (C-4'), 158.0 (C-6'), 143.6 (C-3), 124.6 (C-4), 109.7 (C-5'), 46.9 (C-5), 43.6 (C-1), 28.2 (C-2), 21.0 (C-7').

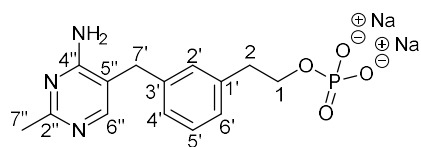
³¹P NMR (162 MHz, D₂O) δ 0.94.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₀H₁₅N₆O₄P: 315.0970; found: 315.0973.

¹H NMR in D₂O:



meta-Phenyl-thiamine monophosphate (*meta*-phenyl-ThMP) **4e**



Prepared from **1e**. White solid (19 mg, 26%).

m.p. 184-186 °C.

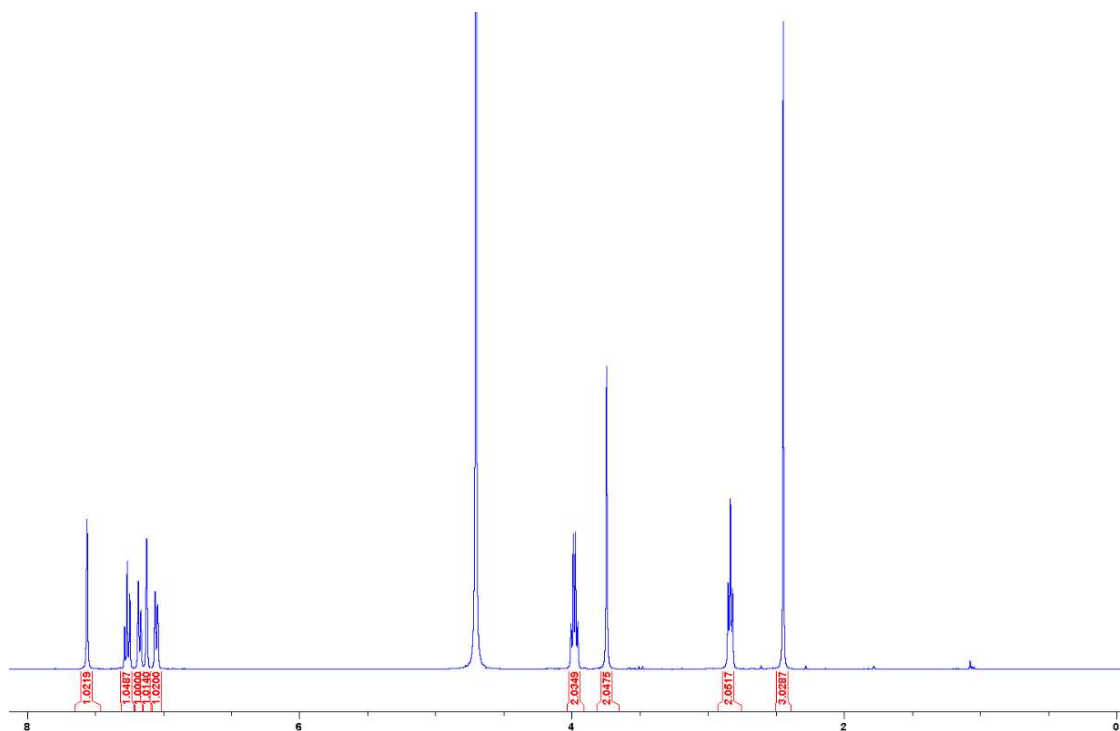
¹H NMR (400 MHz, D₂O) δ 7.56 (s, 1H, H-6''), 7.26 (t, 1H, *J* = 7.6 Hz, H-5'), 7.17 (d, 1H, *J* = 7.6 Hz, H-4'), 7.12 (s, 1H, H-2'), 7.05 (d, 1H, *J* = 7.6 Hz, H-6'), 3.98 (q, 2H, *J* = 6.6 Hz, H-1), 3.74 (s, 2H, H-7'), 2.83 (t, 2H, *J* = 6.6 Hz, H-2), 2.43 (s, 3H, H-7'').

¹³C NMR (100 MHz, D₂O) δ 163.4 (C-2''), 160.8 (C-4''), 140.9 (C-6''), 139.6 (C-3'), 135.4 (C-1'), 129.6 (C-4'), 129.2 (C-6'), 128.0 (C-2'), 126.9 (C-5'), 115.8 (C-5''), 65.9 (C-1), 36.1 (C-2), 32.4 (C-7'), 20.5 (C-7'').

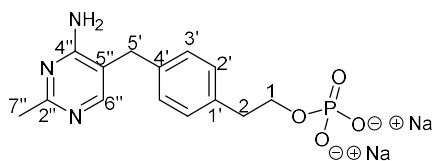
³¹P NMR (162 MHz, D₂O) δ 0.49.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₄H₁₈N₃O₄P: 324.1112; found: 324.1111.

¹H NMR in D₂O:



para-Phenyl-thiamine monophosphate (*para*-phenyl-ThMP) **4f**



Prepared from **1f**. White solid (14 mg, 19%).

m.p. 198-199 °C.

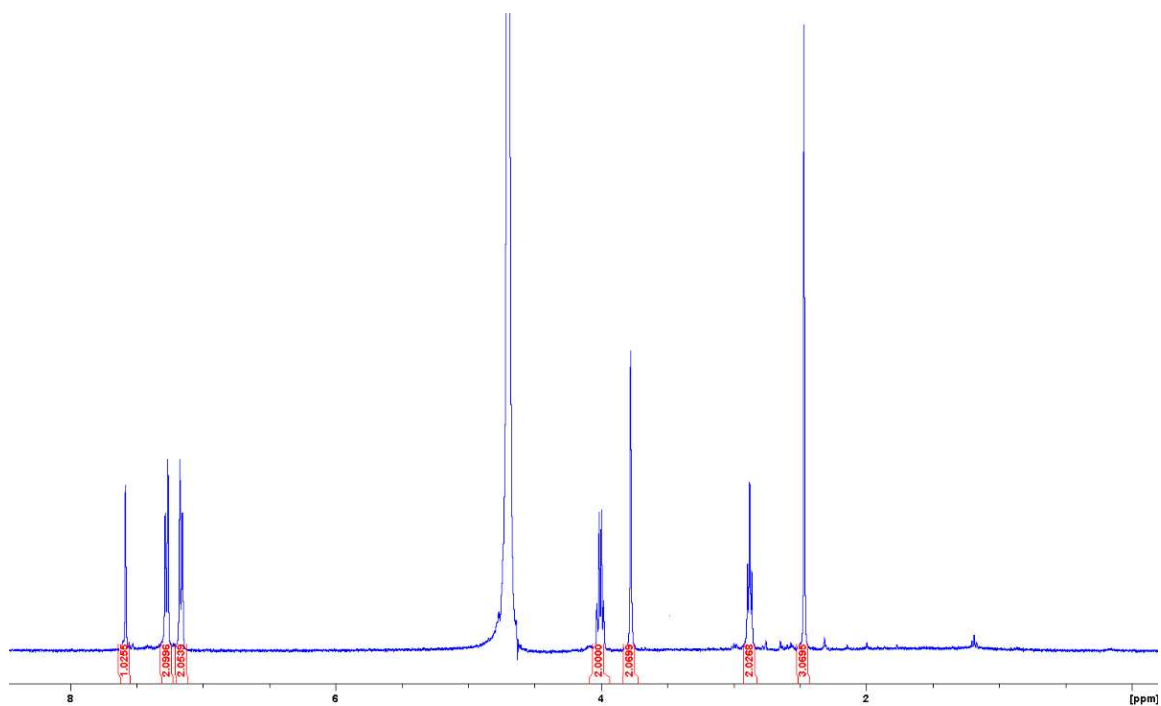
¹H NMR (400 MHz, D₂O) δ 7.59 (s, 1H, H-6''), 7.27 (d, 2H, *J* = 8.0 Hz, H-3'), 7.17 (d, 2H, *J* = 8.0 Hz, H-2'), 4.01 (q, 2H, *J* = 6.8 Hz, H-1), 3.78 (s, 2H, H-5'), 2.88 (t, 2H, *J* = 6.8 Hz, H-2), 2.47 (s, 3H, H-7'').

¹³C NMR (100 MHz, D₂O) δ 163.5 (C-2''), 161.0 (C-4''), 140.7 (C-6''), 137.1 (C-4'), 135.1 (C-1'), 129.7 (C-3'), 128.9 (C-2'), 115.7 (C-5''), 66.0 (C-1), 38.0 (C-2), 32.6 (C-5'), 20.4 (C-7'').

³¹P NMR (162 MHz, D₂O) δ 0.33.

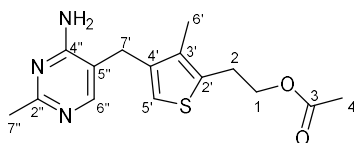
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₄H₁₈N₃O₄P: 324.1112; found: 324.1124.

¹H NMR in D₂O:



Preparation of 5-acyl deaza-ThMP derivatives 8a-h:

2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl}ethyl acetate (deazathiamine acetate) **5**



To a stirred solution of deazathiamine **1a** (132 mg, 0.5 mmol) in dry DMF (2.5 mL, 0.2 M) under nitrogen at 0 °C was added acetic acid (0.038 mL, 0.65 mmol), DCC (310 mg, 1.5 mmol) and DMAP (80 mg, 0.65 mmol). The resultant mixture was stirred at r.t. for 2 days, diluted with DCM, filtered through cotton wool (to remove DCC/DCU), treated with aqueous phosphate buffer (pH 7) (50 mL), and extracted with DCM (100 mL). The organic phase was dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5% MeOH in DCM) to yield deazathiamine acetate **5** as a brown solid (90 mg, 60%).

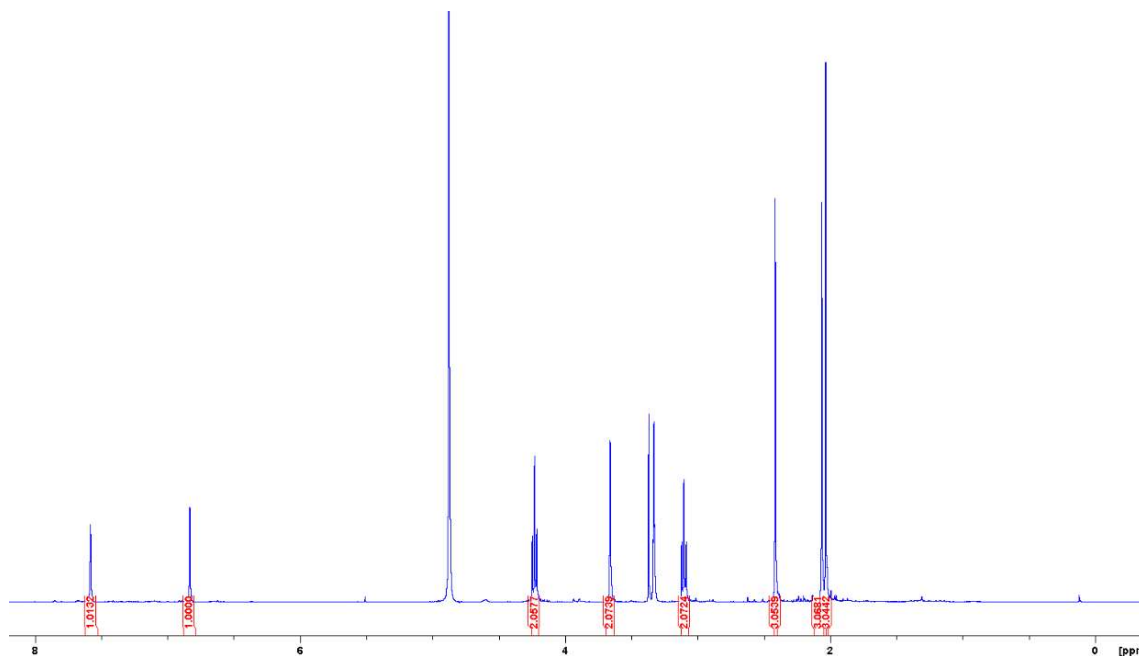
m.p. 117-118 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.58 (s, 1H, H-6''), 6.84 (s, 1H, H-5''), 4.23 (t, 2H, *J* = 6.7 Hz, H-1), 3.66 (s, 2H, H-7'), 3.10 (t, 2H, *J* = 6.7 Hz, H-2), 2.41 (s, 3H, H-7''), 2.06 (s, 3H, H-6'), 2.03 (s, 3H, H-4).

¹³C NMR (100 MHz, CD₃OD) δ 171.3 (C-3), 164.8 (C-2''), 162.3 (C-4''), 153.0 (C-6''), 137.2 (C-4'), 134.3 (C-3'), 133.1 (C-2'), 119.3 (C-5'), 112.8 (C-5''), 64.2 (C-1), 27.5 (C-7'), 27.3 (C-2), 23.3 (C-7''), 19.4 (C-4), 10.8 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₅H₁₉N₃O₂S: 306.1276; found: 306.1282.

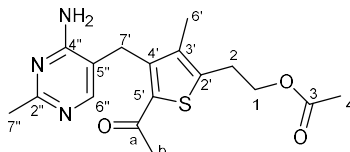
¹H NMR in CD₃OD:



General procedure for preparation of **6a-h**:

To a stirred solution of deazathiamine acetate **5** (0.3 mmol, 1 equiv.) and aluminium chloride (5 equiv.) in dry DCM (0.1 M) under nitrogen at 0 °C was added acyl chloride (8 equiv.) dropwise. The reaction mixture was stirred at 0 °C for 30 min and then at r.t. for 3 h, quenched with cold aqueous phosphate buffer (pH 7) (50 mL) at 0 °C, extracted with DCM (100 mL), washed with aqueous phosphate buffer (pH 7) (50 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5-10% MeOH in DCM) to yield ester **6a-h** as a solid.

2-[5-Acetyl-4-[(4-amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl]ethyl acetate **6a**



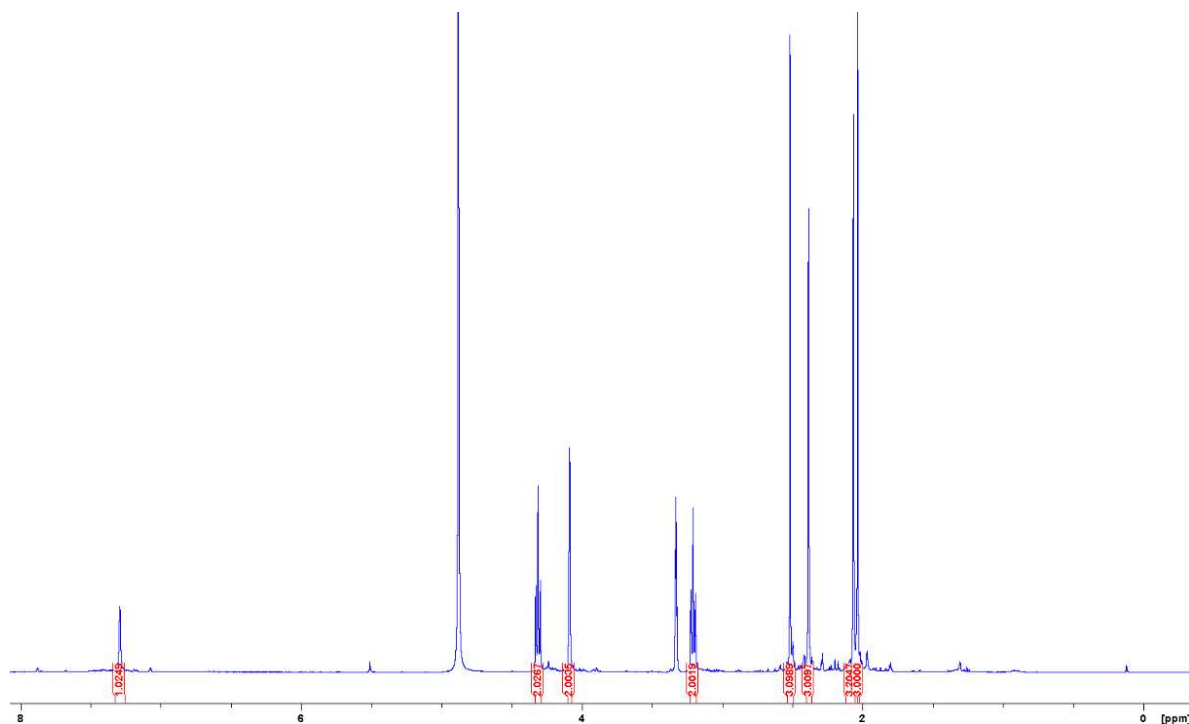
Prepared from acetyl chloride. Yellow solid (39%).

¹H NMR (400 MHz, CD₃OD) δ 7.30 (s, 1H, H-6''), 4.31 (t, 2H, *J* = 6.4 Hz, H-1), 4.09 (s, 2H, H-7'), 3.21 (t, 2H, *J* = 6.4 Hz, H-2), 2.51 (s, 3H, H-b), 2.38 (s, 3H, H-7''), 2.06 (s, 3H, H-6'), 2.03 (s, 3H, H-4).

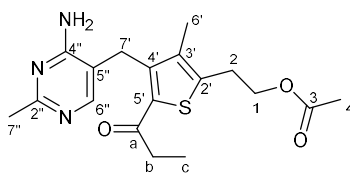
¹³C NMR (100 MHz, CD₃OD) δ 192.0 (C-a), 171.1 (C-3), 164.5 (C-2''), 162.3 (C-4''), 151.1 (C-6''), 143.2 (C-4'), 141.9 (C-2'), 137.5 (C-3'), 134.2 (C-5'), 112.3 (C-5''), 63.5 (C-1), 28.3 (C-b), 27.5 (C-2), 26.1 (C-7'), 23.3 (C-7''), 19.4 (C-4), 10.9 (C-6').

ESI-MS *m/z*: 348.14 [M+H⁺].

¹H NMR in CD₃OD:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-propanoylthiophen-2-yl}ethyl acetate **6b**



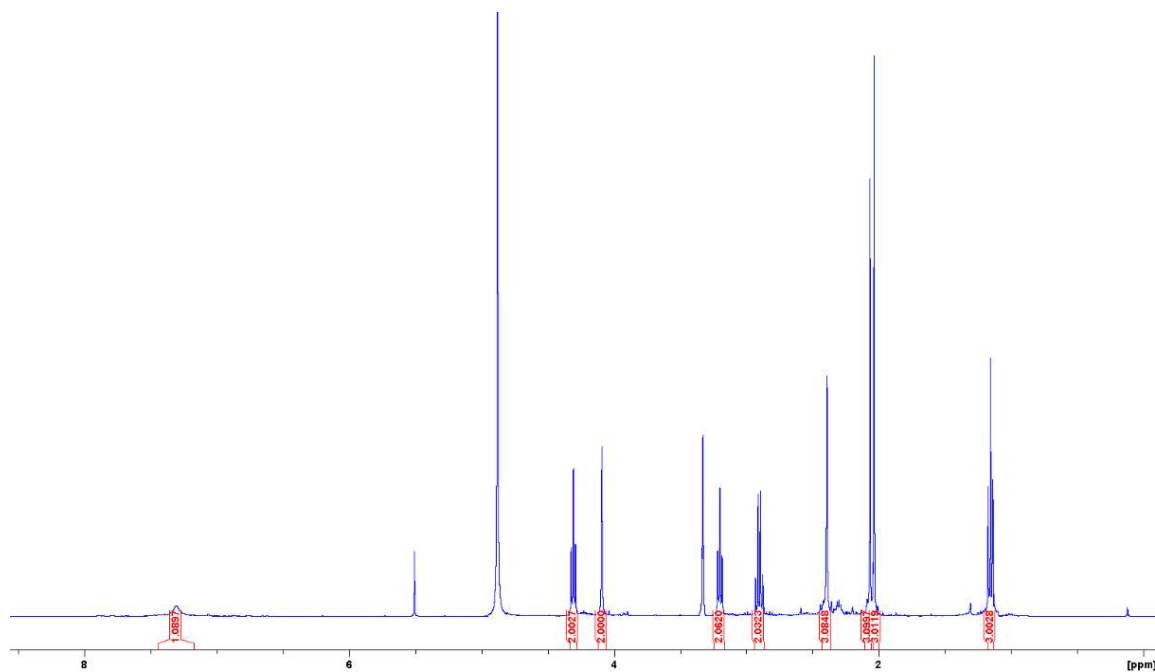
Prepared from propionyl chloride. Yellow solid (46%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.31 (s, 1H, H-6''), 4.30 (t, 2H, $J = 6.6$ Hz, H-1), 4.09 (s, 2H, H-7'), 3.21 (t, 2H, $J = 6.6$ Hz, H-2), 2.90 (q, 2H, $J = 7.3$ Hz, H-b), 2.39 (s, 3H, H-7''), 2.07 (s, 3H, H-6'), 2.03 (s, 3H, H-4), 1.15 (t, 3H, $J = 7.3$ Hz, H-c).

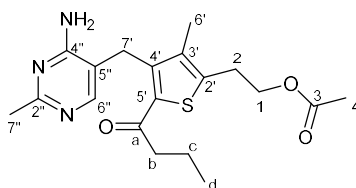
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 195.3 (C-a), 171.1 (C-3), 163.5 (C-2''), 162.5 (C-4''), 150.8 (C-6''), 143.0 (C-4'), 141.1 (C-2'), 137.4 (C-3'), 134.0 (C-5'), 112.1 (C-5''), 63.5 (C-1), 34.7 (C-b), 27.4 (C-2), 26.2 (C-7'), 23.0 (C-7''), 19.3 (C-4), 10.9 (C-6'), 7.5 (C-c).

ESI-MS m/z : 362.15 $[\text{M}+\text{H}^+]$.

$^1\text{H NMR}$ in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-butanoyl-3-methylthiophen-2-yl}ethyl acetate **6c**



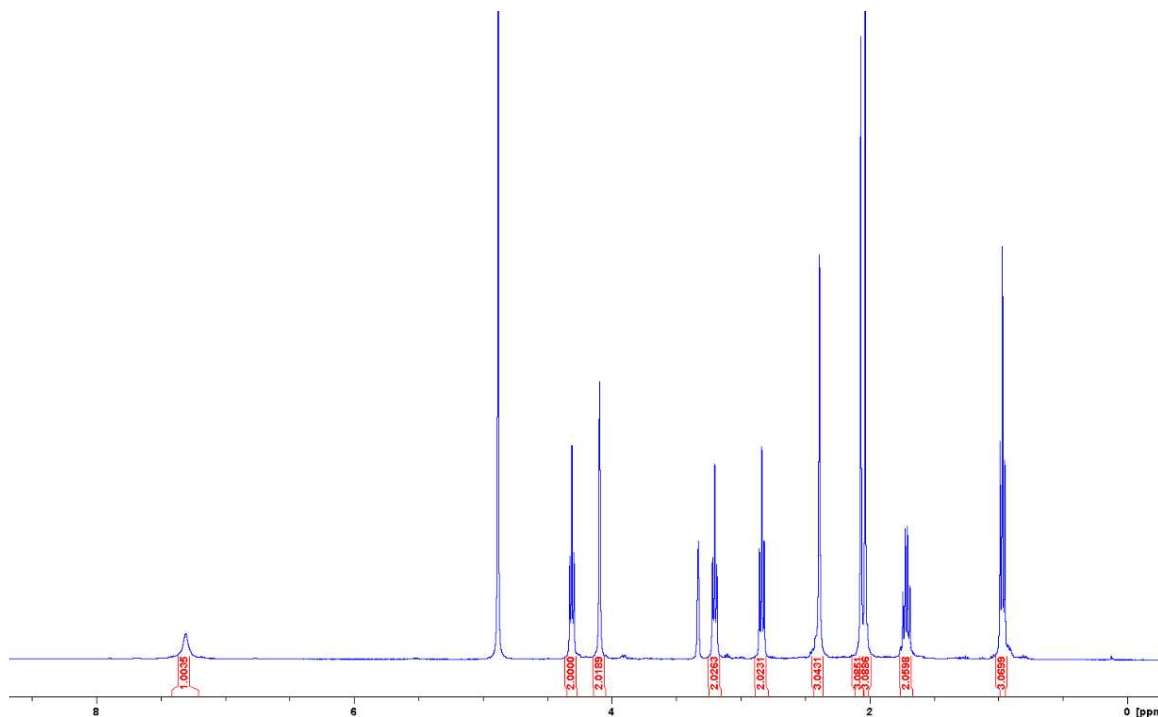
Prepared from butyryl chloride. Yellow solid (41%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.31 (s, 1H, H-6''), 4.31 (t, 2H, $J = 6.5$ Hz, H-1), 4.09 (s, 2H, H-7'), 3.20 (t, 2H, $J = 6.5$ Hz, H-2), 2.84 (t, 2H, $J = 7.2$ Hz, H-b), 2.39 (s, 3H, H-7''), 2.07 (s, 3H, H-6'), 2.04 (s, 3H, H-4), 1.71 (sext, 2H, $J = 7.2$ Hz, H-c), 0.97 (t, 3H, $J = 7.2$ Hz, H-d).

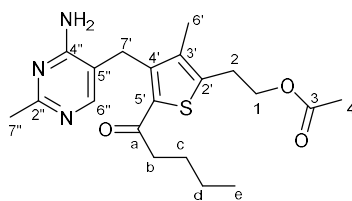
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 195.0 (C-a), 171.1 (C-3), 164.3 (C-2''), 162.5 (C-4''), 150.9 (C-6''), 143.1 (C-4'), 141.4 (C-2'), 137.4 (C-3'), 134.2 (C-5'), 112.1 (C-5''), 63.4 (C-1), 43.5 (C-b), 27.5 (C-2), 26.2 (C-7'), 23.1 (C-7''), 19.4 (C-4), 18.1 (C-c), 12.7 (C-d), 10.8 (C-6').

ESI-MS m/z : 376.17 [$\text{M}+\text{H}^+$].

$^1\text{H NMR}$ in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-pentanoylthiophen-2-yl}ethyl acetate **6d**



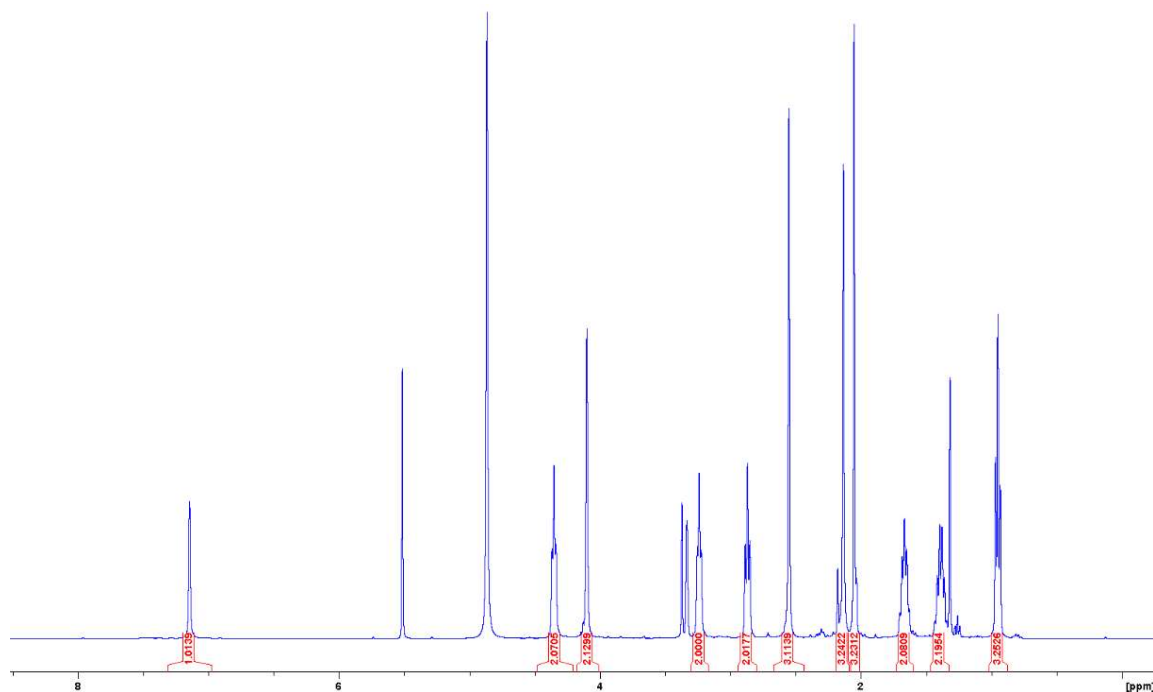
Prepared from valeroyl chloride. Yellow solid (83%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.14 (s, 1H, H-6''), 4.35 (t, 2H, $J = 6.0$ Hz, H-1), 4.10 (s, 2H, H-7'), 3.22 (t, 2H, $J = 6.0$ Hz, H-2), 2.86 (d, 2H, $J = 7.4$ Hz, H-b), 2.55 (s, 3H, H-7''), 2.13 (s, 3H, H-6'), 2.05 (s, 3H, H-4), 1.67 (m, 2H, H-c), 1.39 (m, 2H, H-d), 0.97 (t, 3H, $J = 7.2$ Hz, H-e).

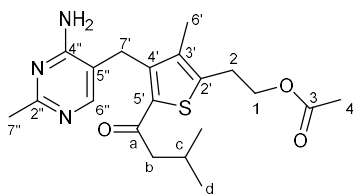
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 194.5 (C-a), 171.3 (C-3), 164.6 (C-2''), 160.2 (C-4''), 141.8 (C-6''), 140.9 (C-4'), 138.6 (C-2'), 137.3 (C-3'), 134.5 (C-5'), 114.1 (C-5''), 63.3 (C-1), 53.6 (C-a), 41.3 (C-b), 27.7 (C-2), 26.7 (C-c), 26.1 (C-7'), 22.0 (C-d), 20.2 (C-7''), 19.5 (C-4), 12.9 (C-e), 10.8 (C-6').

ESI-MS m/z : 390.18 [$\text{M}+\text{H}^+$].

$^1\text{H NMR}$ in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-(3-methylbutanoyl)thiophen-2-yl}ethyl acetate **6e**



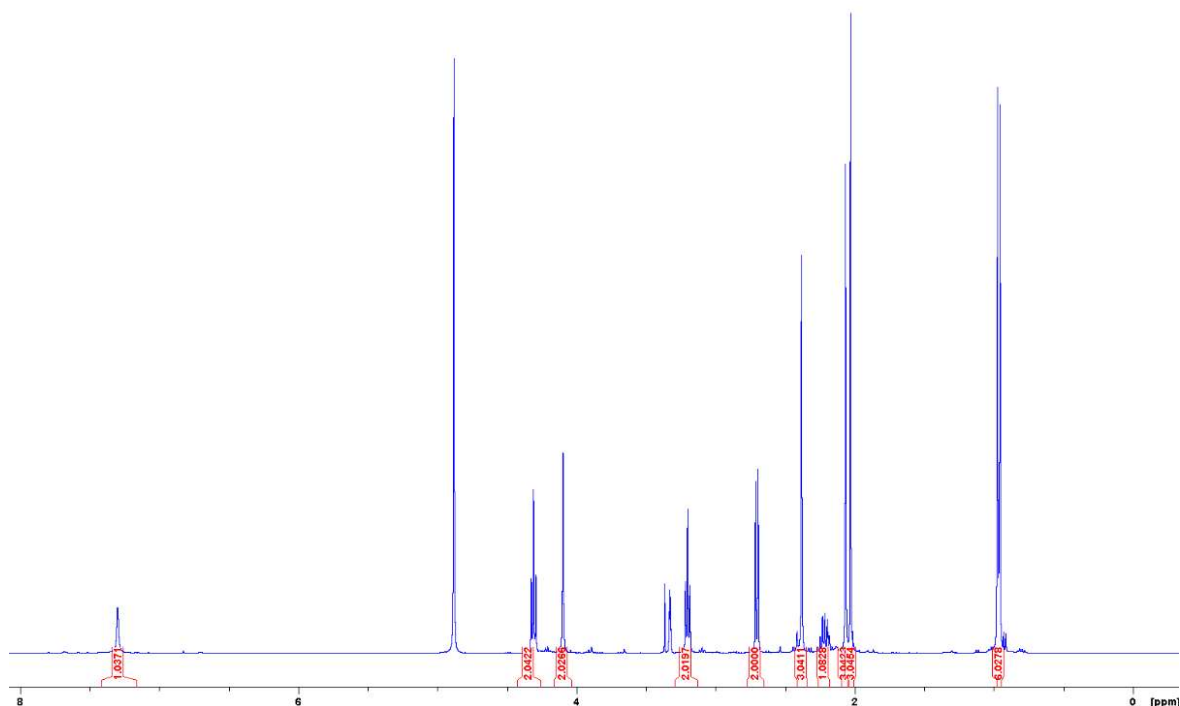
Prepared from isovaleryl chloride. Yellow solid (60%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.33 (s, 1H, H-6''), 4.31 (t, 2H, $J = 6.4$ Hz, H-1), 4.10 (s, 2H, H-7'), 3.20 (t, 2H, $J = 6.4$ Hz, H-2), 2.71 (d, 2H, $J = 7.0$ Hz, H-b), 2.38 (s, 3H, H-7''), 2.20 (m, 1H, H-c), 2.07 (s, 3H, H-6'), 2.03 (s, 3H, H-4), 0.97 (d, 6H, $J = 6.6$ Hz, H-d).

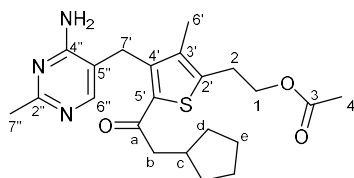
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 194.6 (C-a), 171.0 (C-3), 164.5 (C-2''), 162.5 (C-4''), 151.1 (C-6''), 143.2 (C-4'), 141.4 (C-2'), 137.5 (C-3'), 134.5 (C-5'), 112.1 (C-5''), 63.5 (C-1), 50.6 (C-a), 27.5 (C-2), 26.2 (C-7'), 25.9 (C-c), 23.2 (C-7''), 21.5 (C-d), 19.4 (C-4), 10.9 (C-6').

ESI-MS m/z : 390.18 [$\text{M}+\text{H}^+$].

$^1\text{H NMR}$ in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-cyclopentylacetyl)-3-methylthiophen-2-yl}ethyl acetate **6f**



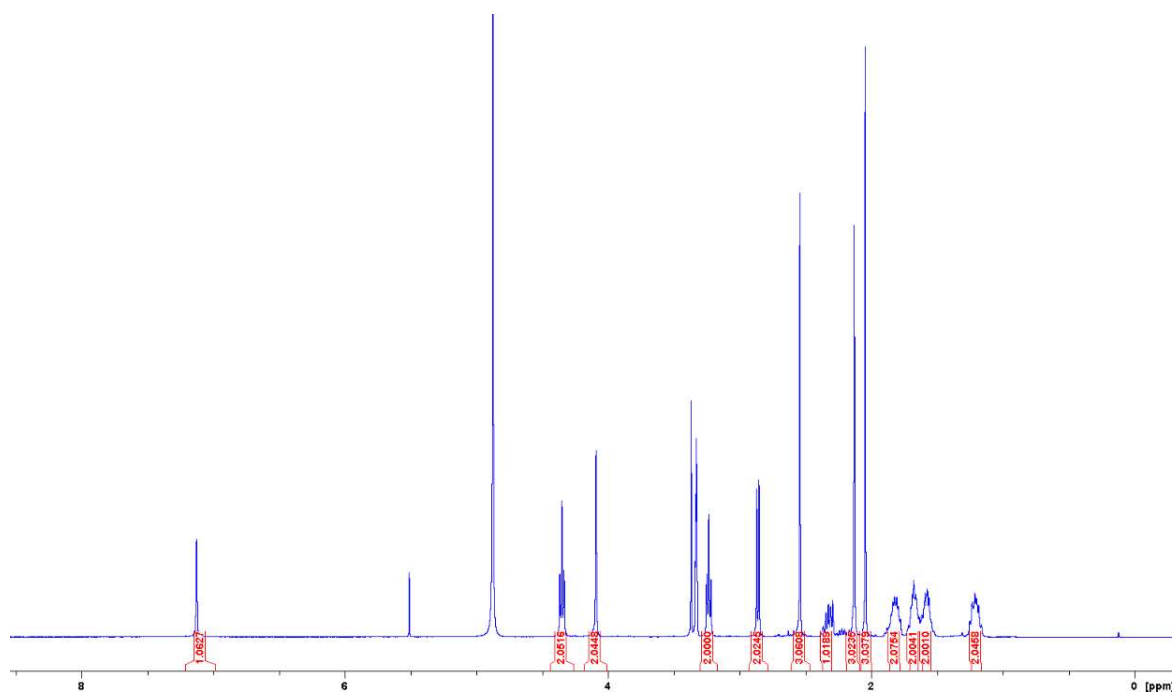
Prepared from cyclopentylacetyl chloride. Yellow solid (60%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.13 (s, 1H, H-6''), 4.35 (t, 2H, $J = 6.4$ Hz, H-1), 4.10 (s, 2H, H-7'), 3.23 (t, 2H, $J = 6.4$ Hz, H-2), 2.86 (d, 2H, $J = 7.0$ Hz, H-b), 2.54 (s, 3H, H-7''), 2.33 (m, 1H, H-c), 2.13 (s, 3H, H-6'), 2.04 (s, 3H, H-4), 1.82 (m, 2H, H-d), 1.67 (m, 2H, H-e), 1.58 (m, 2H, H-e), 1.21 (m, 2H, H-d).

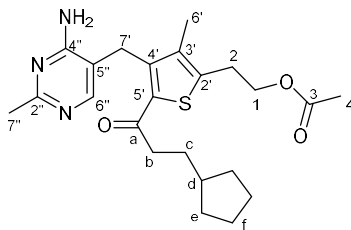
$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 194.6 (C-a), 171.1 (C-3), 164.5 (C-2''), 160.2 (C-4''), 141.9 (C-6''), 140.9 (C-4'), 138.5 (C-2'), 137.3 (C-3'), 134.7 (C-5'), 114.2 (C-5''), 63.2 (C-1), 47.8 (C-b), 36.8 (C-c), 32.1 (C-d), 27.7 (C-2), 26.1 (C-7'), 24.5 (C-e), 20.1 (C-7''), 19.4 (C-4), 10.7 (C-6').

ESI-MS m/z : 416.20 $[\text{M}+\text{H}^+]$.

$^1\text{H NMR}$ in CD_3OD :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(3-cyclopentylpropanoyl)-3-methylthiophen-2-yl}ethyl acetate **6g**



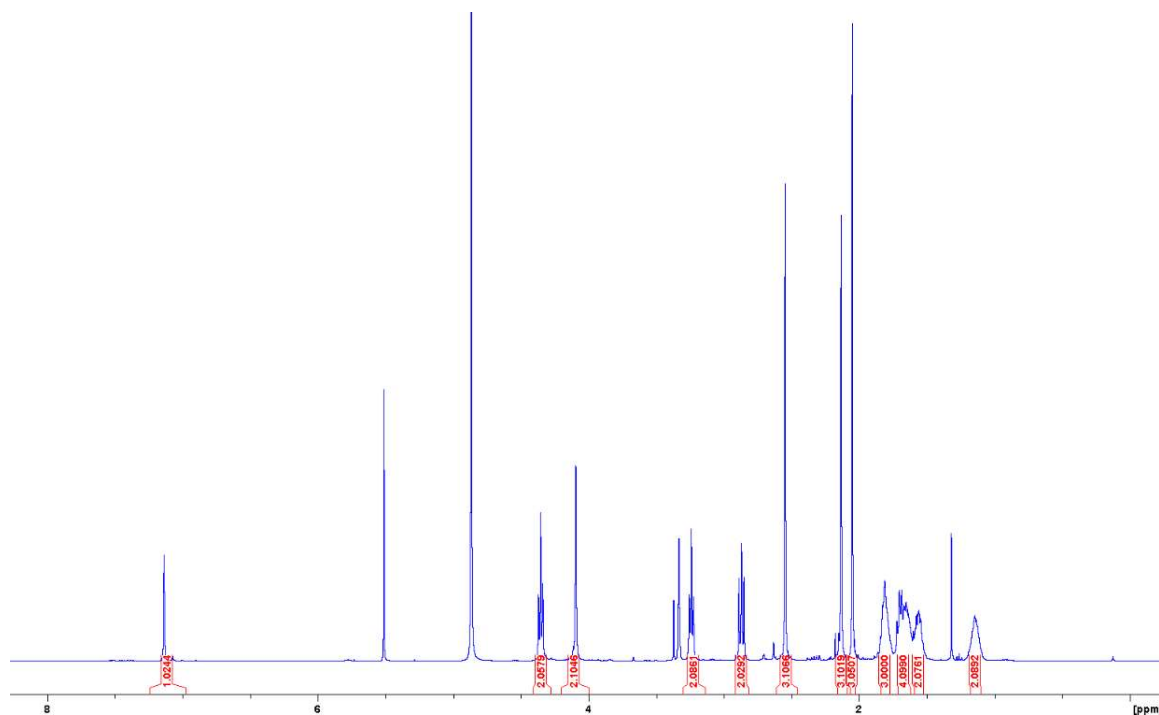
Prepared from 3-cyclopentylpropanoyl chloride. Yellow solid (79%).

¹H NMR (400 MHz, CD₃OD) δ 7.16 (s, 1H, H-6''), 4.35 (t, 2H, *J* = 6.7 Hz, H-1), 4.10 (s, 2H, H-7'), 3.24 (t, 2H, *J* = 6.7 Hz, H-2), 2.87 (t, 2H, *J* = 7.3 Hz, H-b), 2.55 (s, 3H, H-7''), 2.13 (s, 3H, H-6'), 2.05 (s, 3H, H-4), 1.82 (m, 3H, H-d and H-e), 1.68 (m, 4H, H-c and H-f), 1.56 (m, 2H, H-f), 1.15 (m, 2H, H-e).

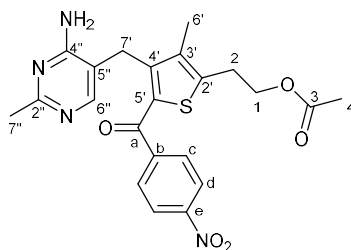
¹³C NMR (100 MHz, CD₃OD) δ 194.6 (C-a), 171.0 (C-3), 164.5 (C-2''), 160.1 (C-4''), 141.9 (C-6''), 140.9 (C-4'), 138.5 (C-2'), 137.3 (C-3'), 134.5 (C-5'), 114.2 (C-5''), 63.4 (C-1), 40.9 (C-b), 39.7 (C-d), 32.2 (C-e), 31.1 (C-c), 27.7 (C-2), 26.1 (C-7'), 24.7 (C-f), 20.1 (C-7''), 19.4 (C-4), 10.8 (C-6').

ESI-MS *m/z*: 430.21 [M+H⁺].

¹H NMR in CD₃OD:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-(4-nitrobenzoyl)thiophen-2-yl}ethyl acetate **6h**



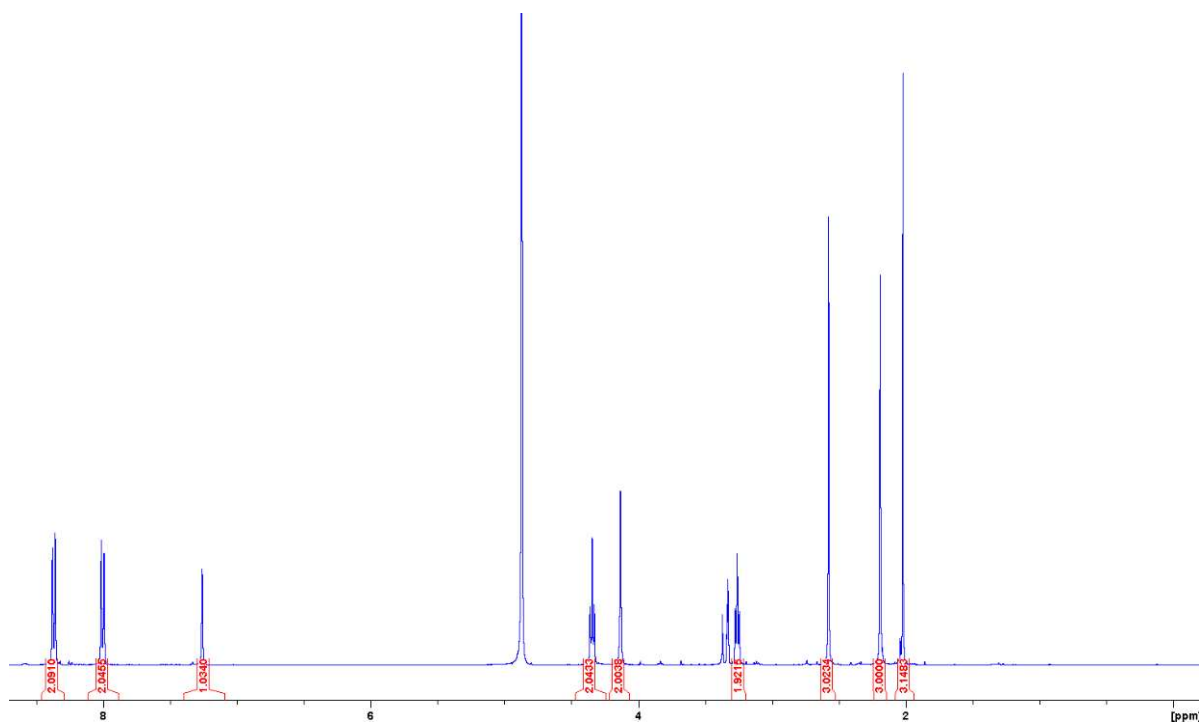
Prepared from 4-nitrobenzoyl chloride. Yellow solid (45%).

$^1\text{H NMR}$ (400 MHz, CD_3OD) δ 8.37 (d, 2H, $J = 8.7$ Hz, H-d), 8.00 (d, 2H, $J = 8.7$ Hz, H-c), 7.26 (s, 1H, H-6''), 4.34 (t, 2H, $J = 6.5$ Hz, H-1), 4.13 (s, 2H, H-7'), 3.25 (t, 2H, $J = 6.5$ Hz, H-2), 2.38 (s, 3H, H-7''), 2.19 (s, 3H, H-6'), 2.03 (s, 3H, H-4).

$^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 187.1 (C-a), 171.1 (C-3), 164.5 (C-2''), 160.2 (C-4''), 149.8 (C-e), 145.3 (C-2'), 144.8 (C-b), 143.5 (C-4'), 139.0 (C-6''), 137.8 (C-3'), 132.9 (C-5'), 129.6 (C-c), 123.2 (C-d), 113.9 (C-5''), 63.0 (C-1), 27.8 (C-2), 26.2 (C-7'), 20.1 (C-7''), 19.5 (C-4), 10.7 (C-6').

ESI-MS m/z : 455.13 [$\text{M}+\text{H}^+$].

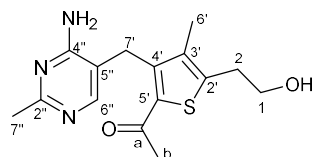
$^1\text{H NMR}$ in CD_3OD :



General procedure for preparation of **7a-h**:

To a stirred solution of ester **6a-h** (0.15 mmol, 1 equiv.) in dry MeOH (0.2 M) under nitrogen was added potassium carbonate (0.3 mmol, 2 equiv.). The resultant mixture was stirred at r.t. for 3 h, concentrated under reduced pressure, diluted with *n*-BuOH (50 mL), washed with sat. aq. NaHCO₃ (15 mL), dried over MgSO₄, filtered, and evaporated under reduced pressure. The residue was purified by silica flash chromatography (5-10% MeOH in DCM) to yield alcohol **7a-h** as a solid.

1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}ethan-1-one (5-acetyl-deazathiamine) **7a**



Prepared from **6a**. White solid (91%).

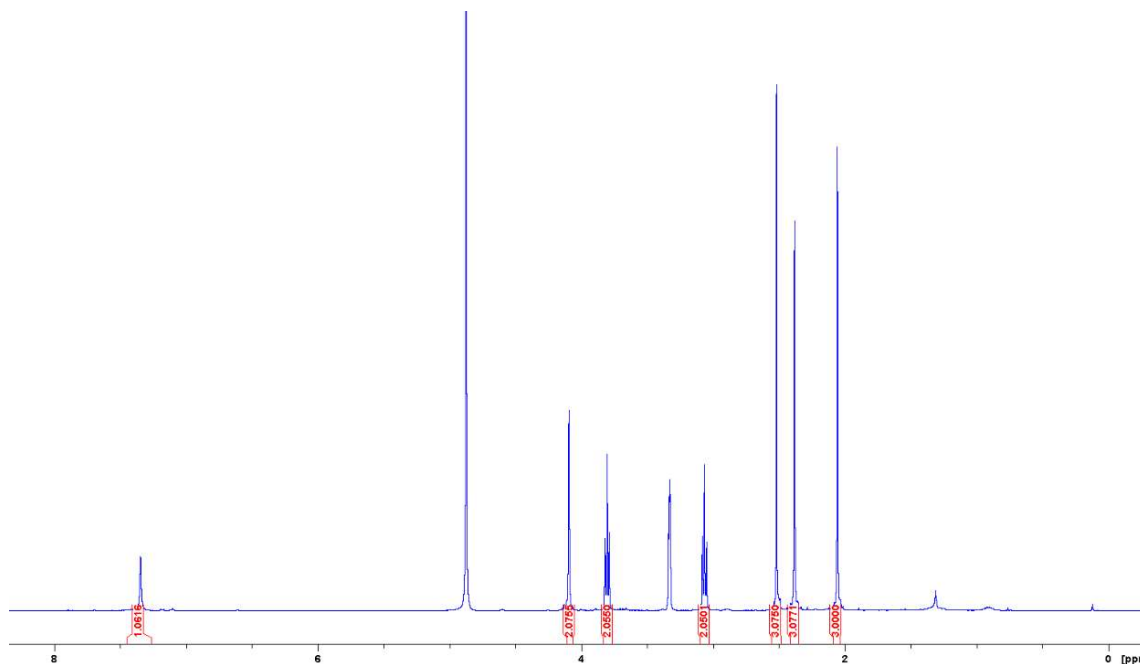
m.p. 204-205 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.35 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.3 Hz, H-1), 3.07 (t, 2H, *J* = 6.3 Hz, H-2), 2.52 (s, 3H, H-b), 2.38 (s, 3H, H-7''), 2.06 (s, 3H, H-6').

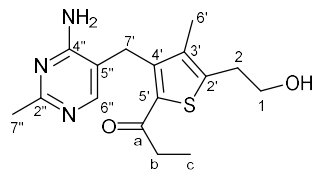
¹³C NMR (100 MHz, CD₃OD) δ 192.3 (C-a), 164.7 (C-2''), 162.5 (C-4''), 151.8 (C-6''), 143.4 (C-4'), 143.2 (C-2'), 137.1 (C-3'), 134.0 (C-5'), 112.0 (C-5''), 61.6 (C-1), 31.4 (C-2), 28.2 (C-b), 26.1 (C-7'), 23.2 (C-7''), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₅H₁₉N₃O₂S: 306.1276; found: 306.1280.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}propan-1-one (5-propanoyl-deazathiamine) **7b**



Prepared from **6b**. White solid (77%).

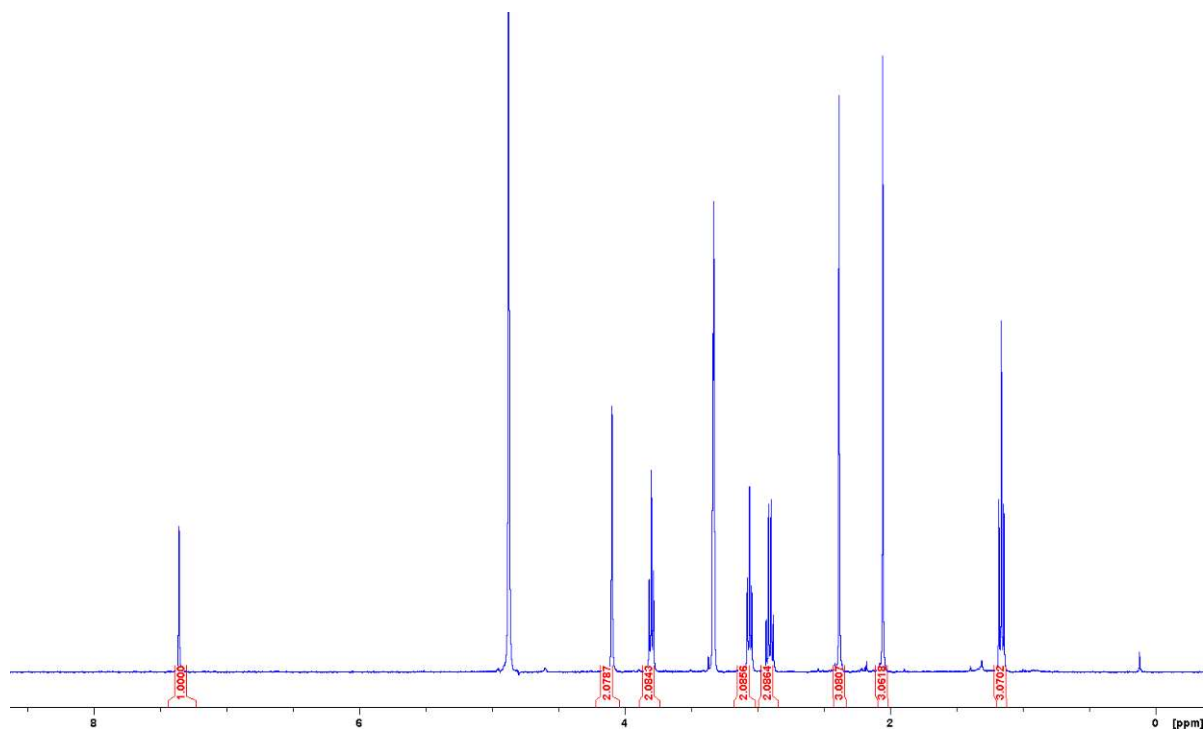
m.p. 211-212 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.36 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.5 Hz, H-1), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.91 (q, 2H, *J* = 7.4 Hz, H-b), 2.38 (s, 3H, H-7''), 2.06 (s, 3H, H-6'), 1.16 (t, 3H, *J* = 7.4 Hz, H-c).

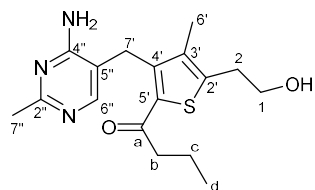
¹³C NMR (100 MHz, CD₃OD) δ 193.3 (C-a), 164.7 (C-2''), 162.3 (C-4''), 151.9 (C-6''), 143.0 (C-4'), 142.8 (C-2'), 137.5 (C-3'), 134.2 (C-5'), 111.9 (C-5''), 61.6 (C-1), 34.7 (C-b), 31.6 (C-2), 26.3 (C-7''), 23.2 (C-7'), 10.9 (C-6'), 7.7 (C-c).

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₆H₂₁N₃O₂S: 320.1432; found: 320.1435.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}butan-1-one (5-butanoyl-deazathiamine) **7c**



Prepared from **6c**. White solid (95%).

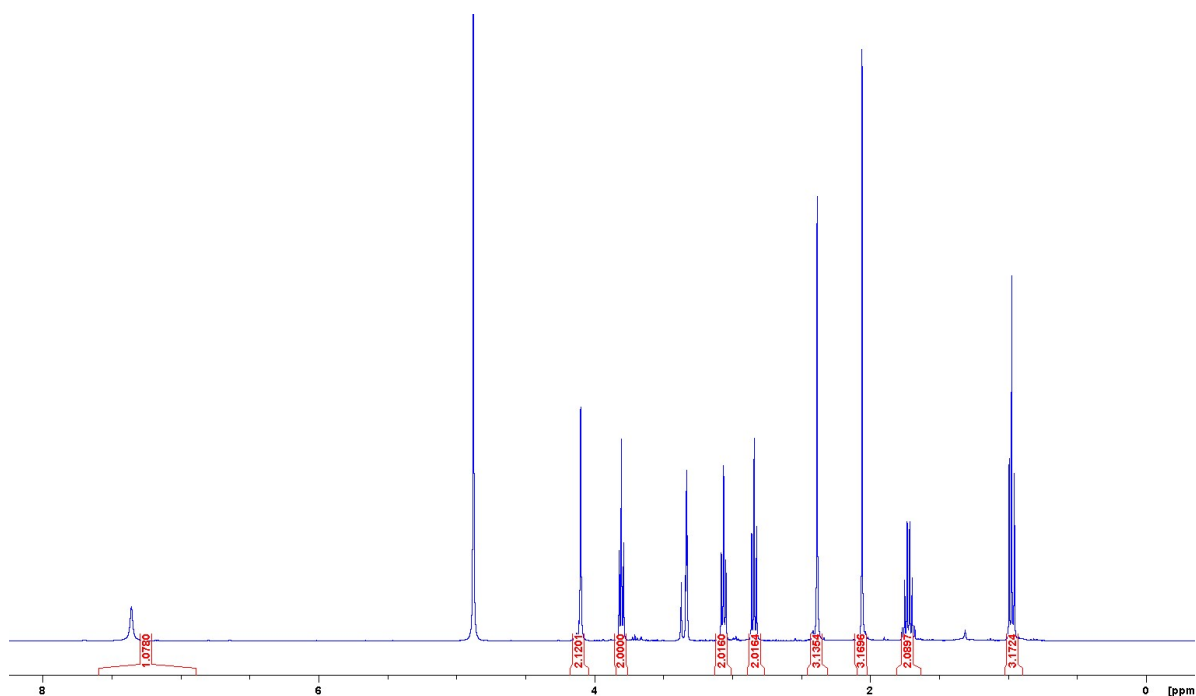
m.p. 214-215 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.35 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.5 Hz, H-1), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.84 (t, 2H, *J* = 7.3 Hz, H-b), 2.38 (s, 3H, H-7''), 2.06 (s, 3H, H-6'), 1.72 (sext, 2H, *J* = 7.2 Hz, H-c), 0.97 (t, 3H, *J* = 7.2 Hz, H-d).

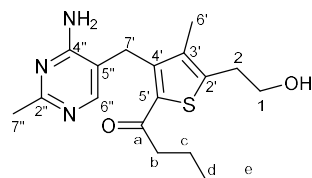
¹³C NMR (100 MHz, CD₃OD) δ 195.2 (C-a), 164.3 (C-2''), 162.4 (C-4''), 151.9 (C-6''), 143.1 (C-4'), 142.8 (C-2'), 137.0 (C-3'), 133.8 (C-5'), 112.0 (C-5''), 61.5 (C-1), 43.5 (C-b), 31.5 (C-2), 26.2 (C-7'), 23.3 (C-7''), 18.1 (C-c), 12.7 (C-d), 10.8 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₇H₂₃N₃O₂S: 334.1589; found: 334.1570.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}pentan-1-one (5-pentanoyl-deazathiamine) **7d**



Prepared from **6d**. White solid (90%).

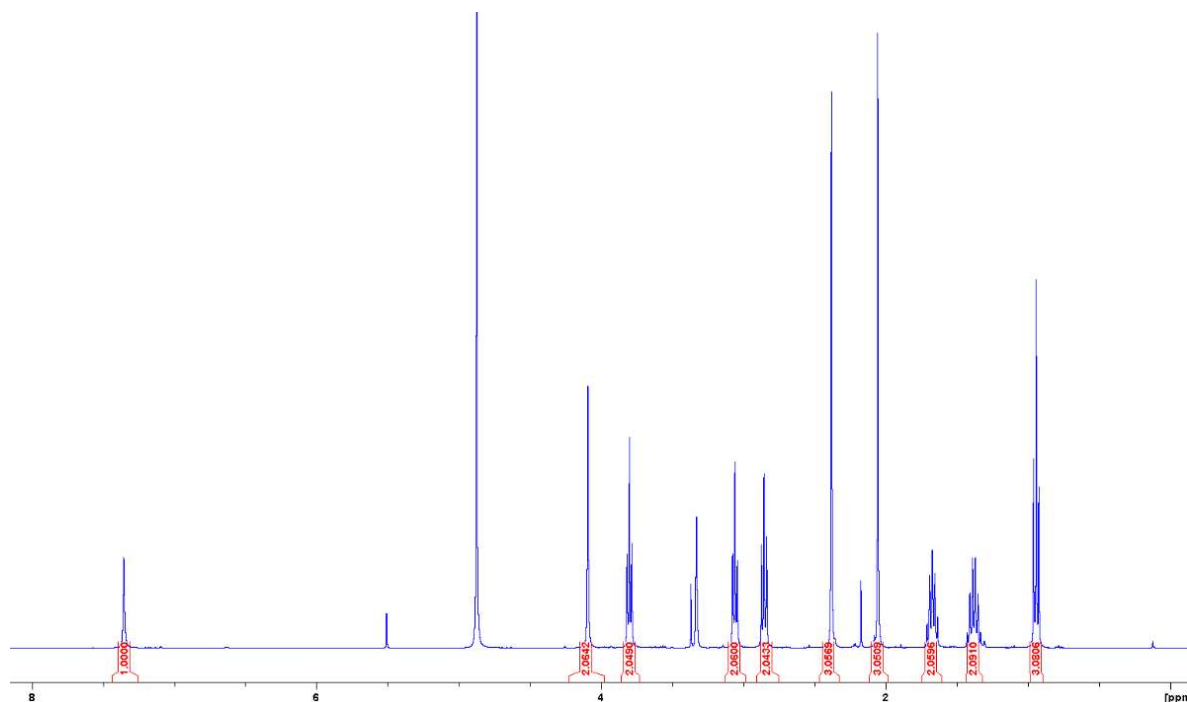
m.p. 216-218 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.35 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.5 Hz, H-1), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.86 (t, 2H, *J* = 7.3 Hz, H-b), 2.38 (s, 3H, H-7''), 2.06 (s, 3H, H-6'), 1.67 (qn, 2H, *J* = 7.3 Hz, H-c), 1.38 (sext, 2H, *J* = 7.3 Hz, H-d), 0.94 (t, 2H, *J* = 7.3 Hz, H-e).

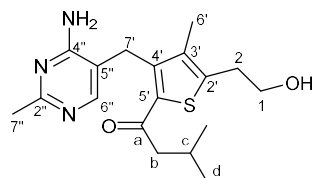
¹³C NMR (100 MHz, CD₃OD) δ 195.7 (C-a), 164.9 (C-2''), 162.4 (C-4''), 151.9 (C-6''), 143.1 (C-4'), 143.0 (C-2'), 137.0 (C-3'), 133.8 (C-5'), 112.0 (C-5''), 61.7 (C-1), 41.5 (C-b), 31.4 (C-2), 27.0 (C-c), 26.3 (C-7'), 23.2 (C-7''), 22.0 (C-d), 12.8 (C-e), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₈H₂₅N₃O₂S: 348.1745; found: 348.1755.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}-3-methylbutan-1-one (5-(3-methylbutanoyl)-deazathiamine) **7e**



Prepared from **6e**. White solid (82%).

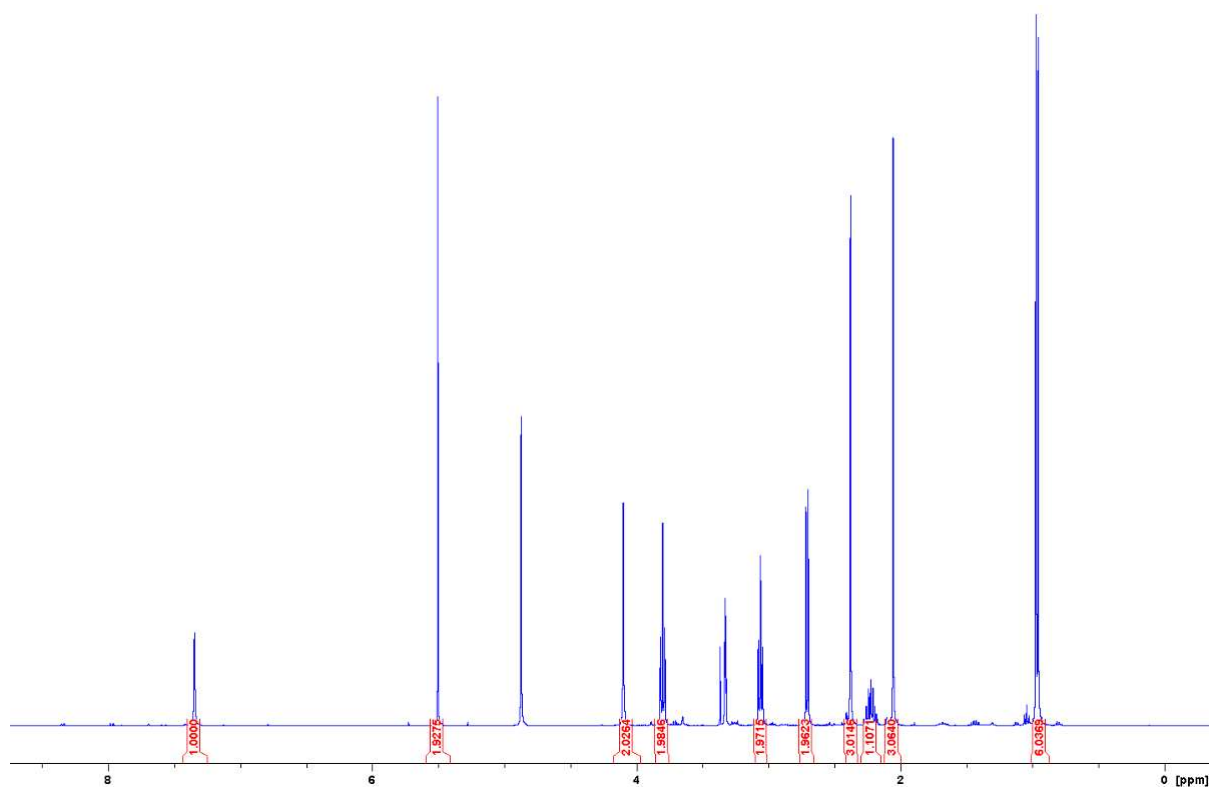
m.p. 202-203 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.35 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.5 Hz, H-1), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.71 (d, 2H, *J* = 7.2 Hz, H-b), 2.38 (s, 3H, H-7''), 2.23 (m, 1H, H-c), 2.06 (s, 3H, H-6'), 0.97 (d, 6H, *J* = 6.6 Hz, H-d).

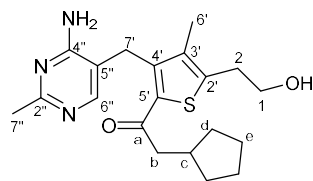
¹³C NMR (100 MHz, CD₃OD) δ 194.7 (C-a), 164.5 (C-2''), 162.5 (C-4''), 151.7 (C-6''), 143.2 (C-4'), 143.1 (C-2'), 137.1 (C-3'), 134.1 (C-5'), 111.9 (C-5''), 61.5 (C-1), 50.6 (C-b), 31.5 (C-2), 26.2 (C-7'), 25.9 (C-c), 23.2 (C-7''), 21.5 (C-d), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₈H₂₅N₃O₂S: 348.1745; found: 348.1749.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}-2-cyclopentylethan-1-one (5-(2-cyclopentylacetyl)-deazathiamine) **7f**



Prepared from **6f**. White solid (69%).

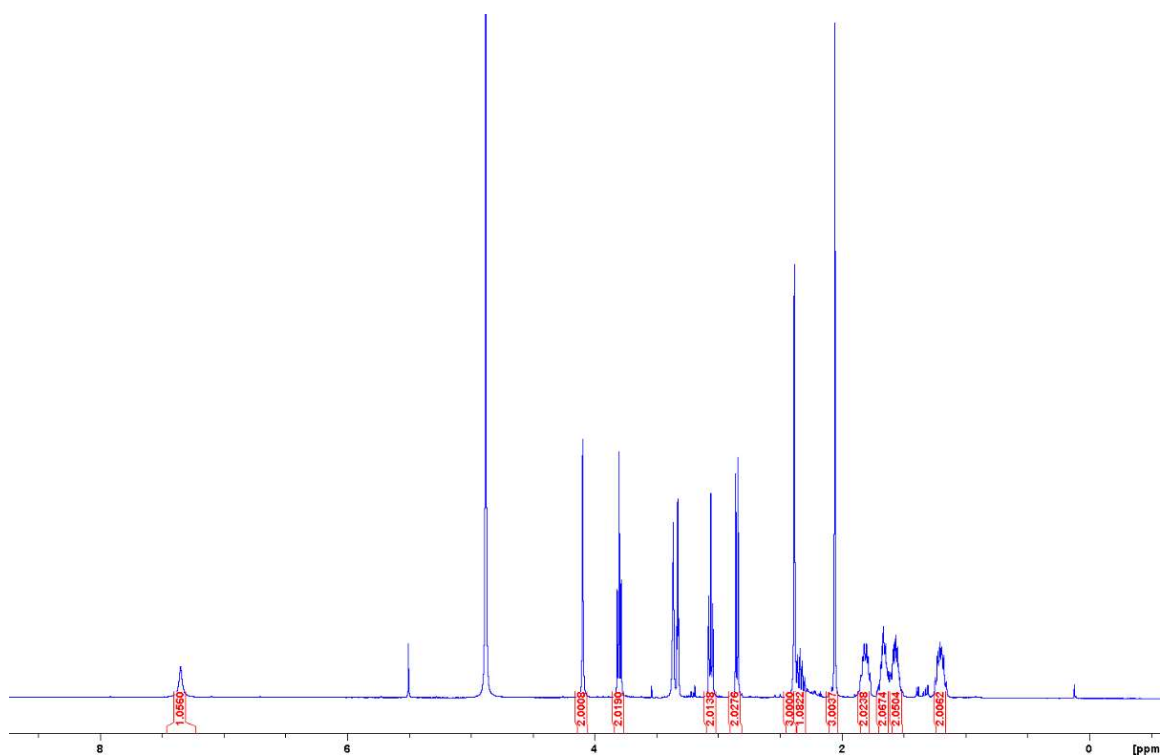
m.p. 233-234 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.35 (s, 1H, H-6''), 4.10 (s, 2H, H-7'), 3.81 (t, 2H, *J* = 6.7 Hz, H-1), 3.07 (t, 2H, *J* = 6.7 Hz, H-2), 2.85 (d, 2H, *J* = 7.2 Hz, H-b), 2.39 (s, 3H, H-7''), 2.34 (m, 1H, H-c), 2.06 (s, 3H, H-6'), 1.81 (m, 2H, H-d), 1.66 (m, 2H, H-e), 1.57 (m, 2H, H-e), 1.21 (m, 2H, H-d).

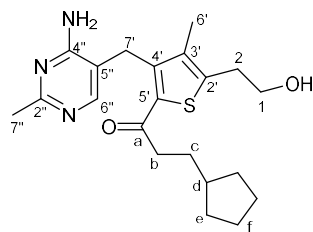
¹³C NMR (100 MHz, CD₃OD) δ 194.9 (C-a), 164.4 (C-2''), 162.4 (C-4''), 151.3 (C-6''), 143.1 (C-4'), 143.1 (C-2'), 137.0 (C-3'), 134.0 (C-5'), 112.1 (C-5''), 61.4 (C-1), 47.7 (C-b), 37.0 (C-c), 32.1 (C-d), 31.5 (C-2), 26.2 (C-7'), 24.4 (C-e), 23.1 (C-7''), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₇N₃O₂S: 374.1902; found: 374.1918.

¹H NMR in CD₃OD:



1-{3-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-hydroxyethyl)-4-methylthiophen-2-yl}-3-cyclopentylpropan-1-one (5-(3-cyclopentylpropanoyl)-deazathiamine) **7g**



Prepared from **6g**. White solid (54%).

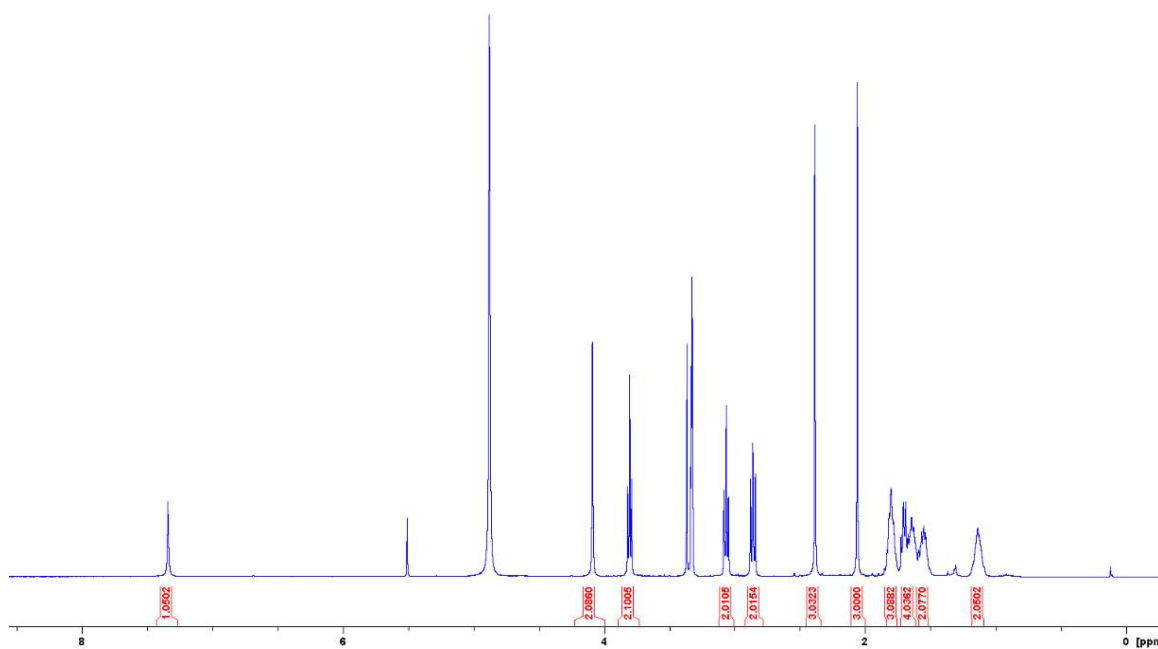
m.p. 241-243 °C.

¹H NMR (400 MHz, CD₃OD) δ 7.33 (s, 1H, H-6''), 4.09 (s, 2H, H-7'), 3.80 (t, 2H, *J* = 6.5 Hz, H-1), 3.06 (t, 2H, *J* = 6.5 Hz, H-2), 2.86 (t, 2H, *J* = 7.5 Hz, H-b), 2.39 (s, 3H, H-7''), 2.07 (s, 3H, H-6'), 1.80 (m, 3H, H-d and H-e), 1.67 (m, 4H, H-c and H-f), 1.56 (m, 2H, H-f), 1.14 (m, 2H, H-e).

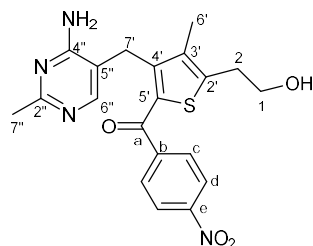
¹³C NMR (100 MHz, CD₃OD) δ 195.2 (C-a), 164.4 (C-2''), 162.4 (C-4''), 151.4 (C-6''), 143.1 (C-4'), 143.0 (C-2'), 137.0 (C-3'), 133.8 (C-5'), 112.1 (C-5''), 61.8 (C-1), 40.9 (C-b), 39.6 (C-d), 32.2 (C-e), 31.5 (C-2), 31.2 (C-c), 26.2 (C-7'), 24.7 (C-f), 23.1 (C-7''), 10.9 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₁H₂₉N₃O₂S: 388.2058; found: 388.2066.

¹H NMR in CD₃OD:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-(4-nitrobenzoyl)thiophen-2-yl}ethan-1-ol
(5-(4-nitrobenzoyl)-deazathiamine) **7h**



Prepared from **6h**. White solid (57%).

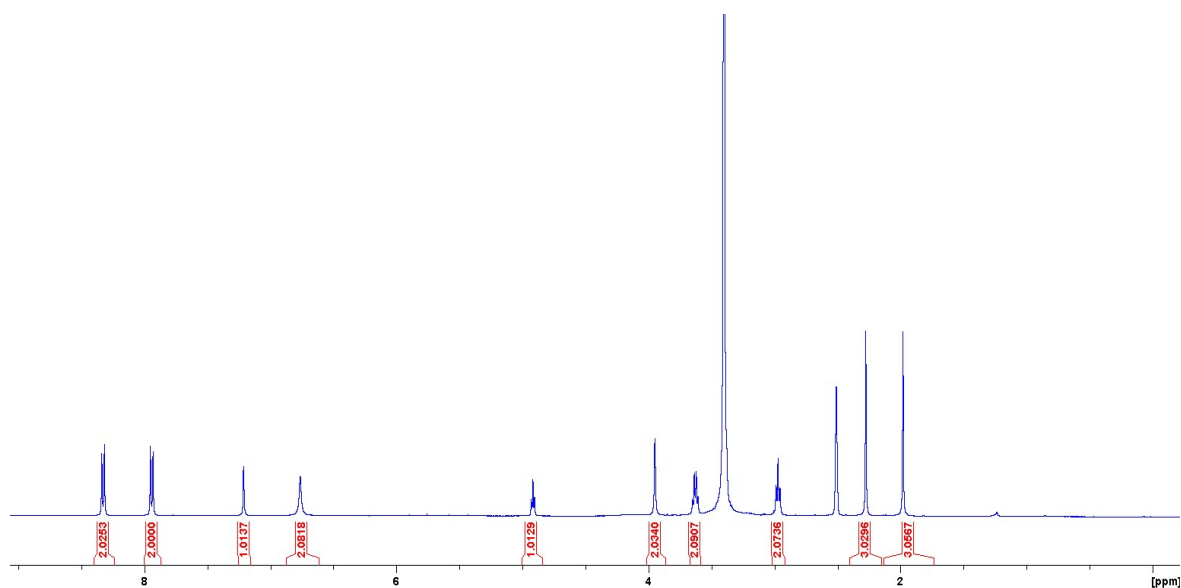
m.p. 222-223 °C.

¹H NMR (400 MHz, CD₃SOCD₃) δ 8.32 (d, 2H, *J* = 8.6 Hz, H-d), 7.94 (d, 2H, *J* = 8.6 Hz, H-c), 7.21 (s, 1H, H-6''), 6.76 (br, NH₂), 4.92 (br, 1H, OH), 3.95 (s, 2H, H-7'), 3.63 (m, 2H, H-1), 2.97 (t, 2H, *J* = 6.2 Hz, H-2), 2.28 (s, 3H, H-7''), 1.98 (s, 3H, H-6').

¹³C NMR (100 MHz, CD₃SOCD₃) δ 187.4 (C-a), 164.7 (C-2''), 162.2 (C-4''), 152.7 (C-6''), 149.8 (C-e), 147.3 (C-4'), 146.6 (C-b), 145.8 (C-2'), 137.5 (C-3'), 132.2 (C-5'), 130.1 (C-c), 123.8 (C-d), 111.6 (C-5''), 61.4 (C-1), 32.3 (C-2), 26.8 (C-7'), 25.3 (C-7''), 12.2 (C-6').

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₀N₄O₄S: 413.1283; found: 413.1288.

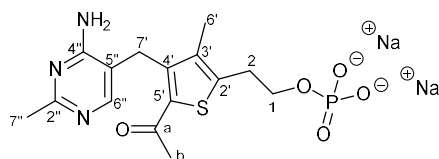
¹H NMR in CD₃SOCD₃:



General procedure for preparation of **8a-h**:

To a flask charged with ester **7a-h** (0.2 mmol, 1 equiv.) under nitrogen at 0 °C was added POCl₃ (1 mL, 0.2M) dropwise. The resultant mixture was stirred at 0 °C for 1 h and then at r.t. for 3 h. The reaction mixture was diluted with cold DCM (15 mL) at 0 °C, quenched with cold sat. aq. NaHCO₃ (10 mL) dropwise (NOTE: this quenching step is highly exothermic!), washed with DCM (2 x 20 mL), and concentrated under reduced pressure. The residue was diluted in minimal amount of water and purified by reverse-phase HPLC to yield monophosphate **8a-h** as a solid.

2-{5-Acetyl-4-[(4-amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl}ethyl dihydrogen phosphate (5-acetyl-deaza-ThMP) **8a**



Prepared from **7a**. White solid (42%).

m.p. 188-189 °C.

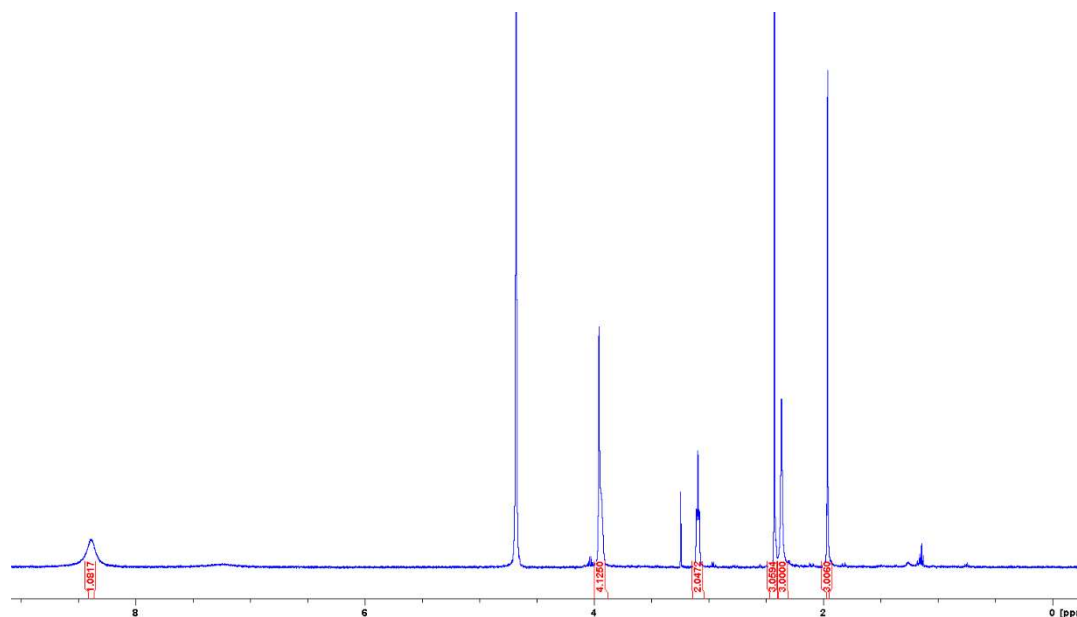
¹H NMR (400 MHz, D₂O) δ 8.39 (s, 1H, H-6''), 3.95 (m, 4H, H-1 and H-7'), 3.09 (t, 2H, *J* = 6.1 Hz, H-2), 2.43 (s, 3H, H-b), 2.37 (s, 3H, H-7''), 1.97 (s, 3H, H-6').

¹³C NMR (100 MHz, D₂O) δ 195.5 (C-a), 171.0 (C-6''), 163.4 (C-2''), 163.2 (C-4''), 144.9 (C-4'), 142.2 (C-2'), 137.7 (C-3'), 134.1 (C-5'), 114.0 (C-5''), 64.5 (C-1), 29.8 (C-2), 28.8 (C-b), 26.2 (C-7'), 21.5 (C-7''), 11.2 (C-6').

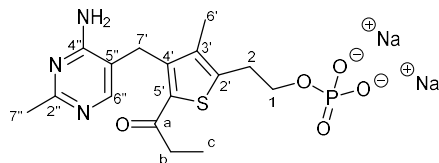
³¹P NMR (162 MHz, D₂O) δ 1.20.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₅H₂₀N₃O₅PS: 386.0940; found: 386.0936.

¹H NMR in D₂O:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-propanoylthiophen-2-yl}ethyl dihydrogen phosphate (5-propanoyl-deaza-ThMP) **8b**



Prepared from **7b**. White solid (10%).

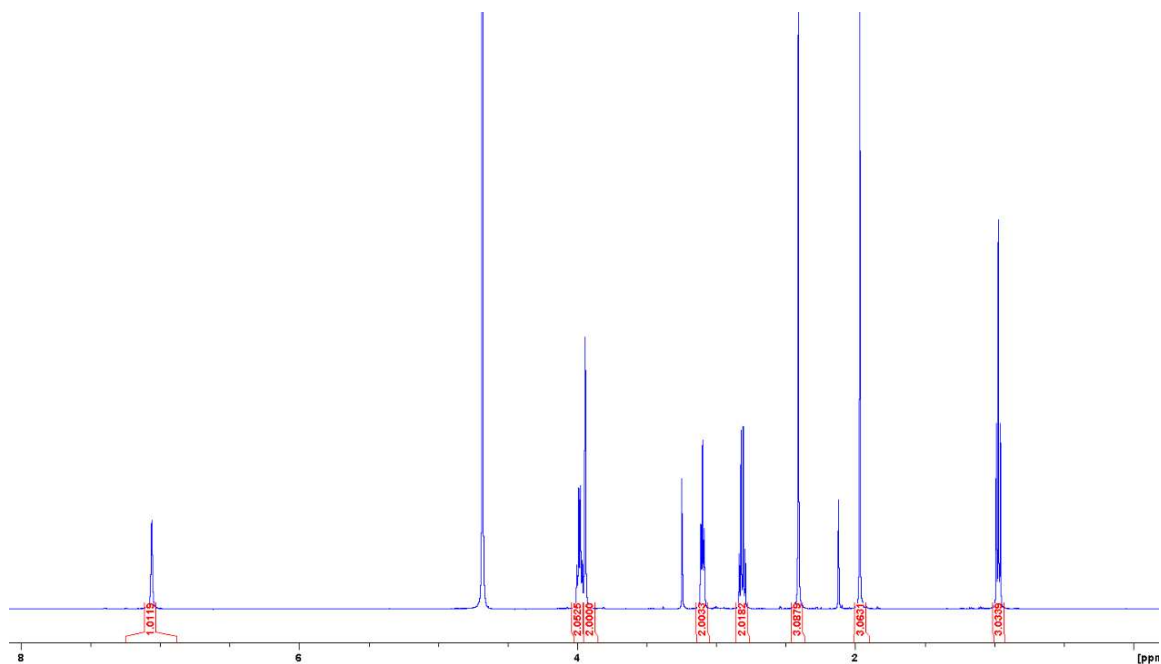
$^1\text{H NMR}$ (400 MHz, D_2O) δ 7.06 (s, 1H, H-6''), 3.98 (q, 2H, $J = 6.4$ Hz, H-1), 3.94 (s, 2H, H-7'), 3.10 (t, 2H, $J = 6.4$ Hz, H-2), 2.81 (q, 2H, $J = 7.4$ Hz, H-b), 2.41 (s, 3H, H-7''), 1.97 (s, 3H, H-6'), 0.97 (t, 2H, $J = 7.4$ Hz, H-c).

$^{13}\text{C NMR}$ (100 MHz, D_2O) δ 198.5 (C-a), 163.8 (C-2''), 160.8 (C-4'), 143.9 (C-4'), 141.4 (C-2'), 139.5 (C-6''), 137.7 (C-3'), 133.8 (C-5'), 114.6 (C-5''), 64.8 (d, $J = 4.6$ Hz, C-1), 34.8 (C-b), 29.5 (C-2), 26.1 (C-7'), 20.6 (C-7''), 11.2 (C-6'), 8.1 (C-c).

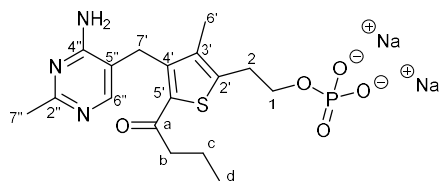
$^{31}\text{P NMR}$ (162 MHz, D_2O) δ -0.9.

HRMS (ESI) m/z : $[\text{M}+\text{H}^+]$ calculated for $\text{C}_{16}\text{H}_{22}\text{N}_3\text{O}_5\text{PS}$: 400.1096; found: 400.1098.

$^1\text{H NMR}$ in D_2O :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-butanoyl-3-methylthiophen-2-yl}ethyl dihydrogen phosphate (5-butanoyl-deaza-ThMP) **8c**



Prepared from **7c**. White solid (15%).

m.p. 204-205 °C.

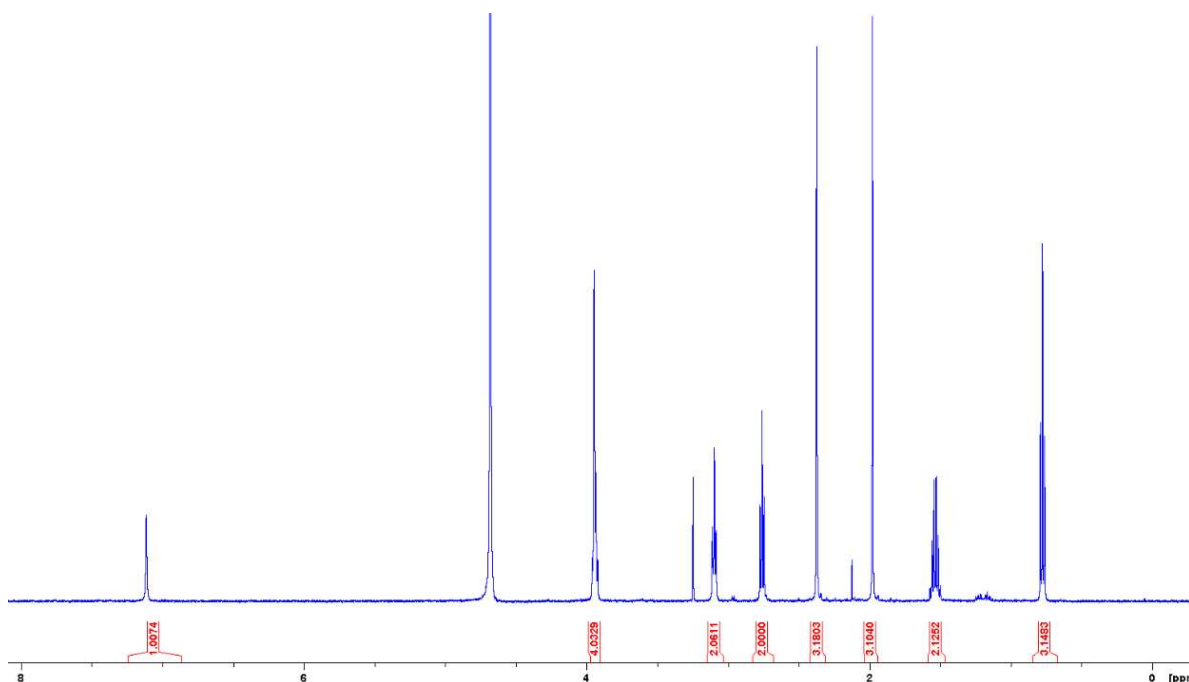
¹H NMR (400 MHz, D₂O) δ 7.11 (s, 1H, H-6''), 3.94 (m, 4H, H-1 and H-7'), 3.10 (t, 2H, *J* = 6.5 Hz, H-2), 2.76 (t, 2H, *J* = 7.4 Hz, H-b), 2.37 (s, 3H, H-7''), 1.98 (s, 3H, H-6'), 1.54 (sext, 2H, *J* = 7.4 Hz, H-c), 0.77 (t, 3H, *J* = 7.4 Hz, H-d).

¹³C NMR (100 MHz, D₂O) δ 198.3 (C-a), 171.1 (C-6''), 163.5 (C-2''), 161.8 (C-4''), 144.4 (C-4'), 142.2 (C-2'), 137.9 (C-3'), 133.9 (C-5'), 114.3 (C-5''), 64.3 (d, *J* = 4.6 Hz, C-1), 43.4 (C-b), 29.8 (C-2), 26.2 (C-7'), 21.3 (C-7''), 18.5 (C-c), 12.8 (C-d), 11.2 (C-6').

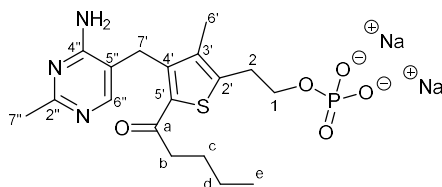
³¹P NMR (162 MHz, D₂O) δ 0.3.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₇H₂₄N₃O₅PS: 414.1252; found: 414.1255.

¹H NMR in D₂O:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-pentanoylthiophen-2-yl}ethyl dihydrogen phosphate (5-pentanoyl-deaza-ThMP) **8d**



Prepared from **7d**. White solid (20%).

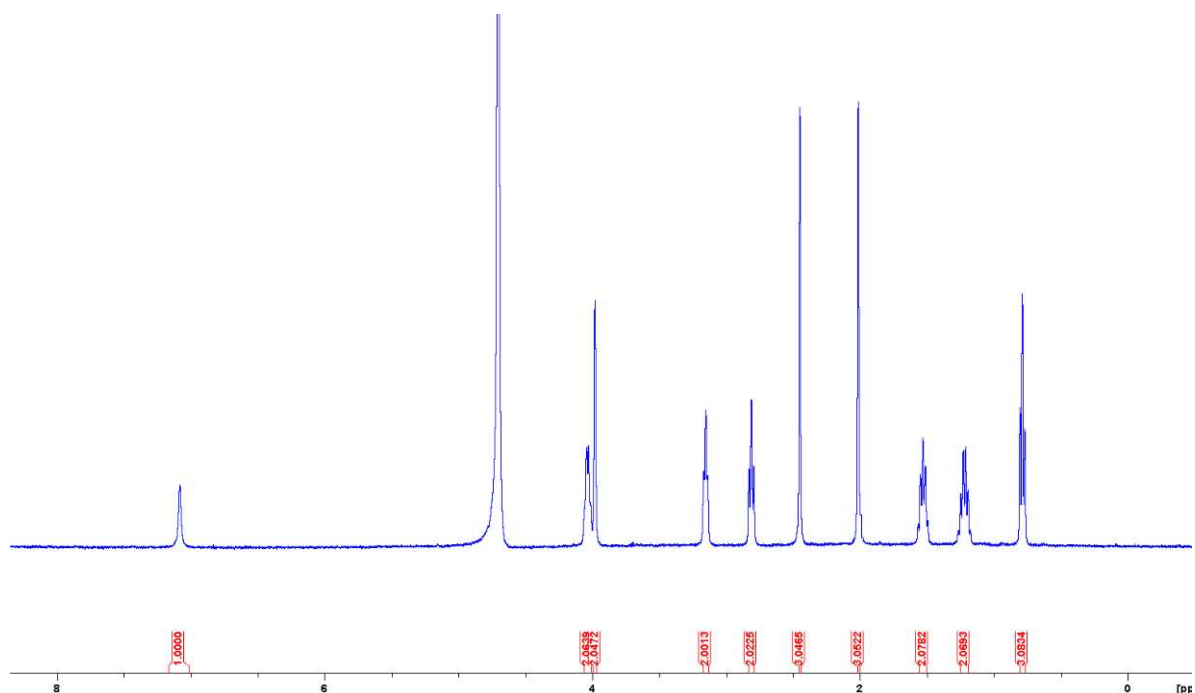
¹H NMR (400 MHz, D₂O) δ 7.08 (s, 1H, H-6''), 4.04 (q, 2H, *J* = 5.8 Hz, H-1), 3.98 (s, 2H, H-7'), 3.15 (t, 2H, *J* = 5.8 Hz, H-2), 2.81 (t, 2H, *J* = 7.1 Hz, H-b), 2.45 (s, 3H, H-7''), 2.01 (s, 3H, H-6'), 1.53 (qn, 2H, *J* = 7.1 Hz, H-c), 1.22 (sext, 2H, *J* = 7.1 Hz, H-d), 0.80 (t, 3H, *J* = 7.1 Hz, H-e).

¹³C NMR (100 MHz, D₂O) δ 198.3 (C-a), 170.9 (C-6''), 163.7 (C-2''), 161.8 (C-4''), 144.3 (C-4'), 141.9 (C-2'), 137.8 (C-3'), 134.0 (C-5'), 115.0 (C-5''), 64.7 (C-1), 41.4 (C-b), 29.6 (C-2), 27.0 (C-c), 26.1 (C-7'), 21.6 (C-d), 20.9 (C-7''), 13.1 (C-e), 11.2 (C-6').

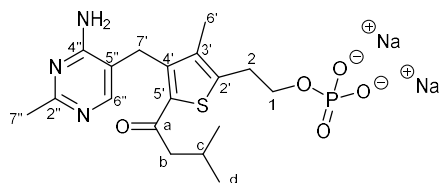
³¹P NMR (162 MHz, D₂O) δ -0.3.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₈H₂₆N₃O₅PS: 428.1409; found: 428.1408.

¹H NMR in D₂O:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-(3-methylbutanoyl)thiophen-2-yl}ethyl dihydrogen phosphate (5-(3-methylbutanoyl)-deaza-ThMP) **8e**



Prepared from **7e**. White solid (23%).

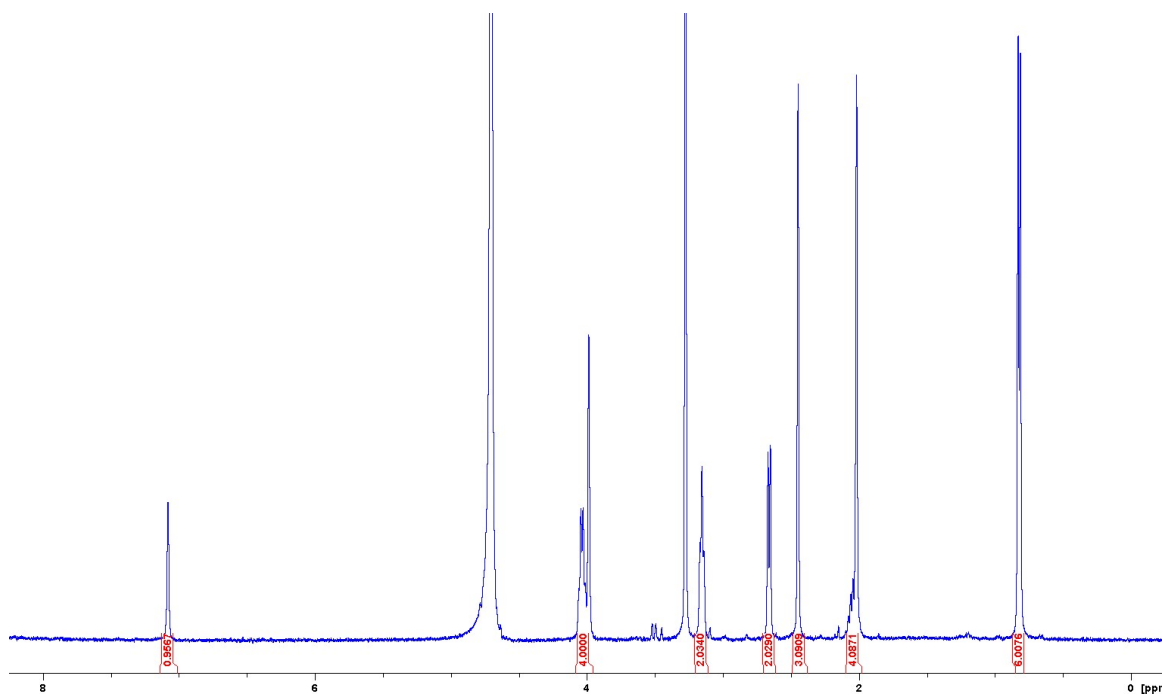
$^1\text{H NMR}$ (400 MHz, D_2O) δ 7.08 (s, 1H, H-6''), 4.04 (q, 2H, $J = 6.0$ Hz, H-1), 3.98 (s, 2H, H-7'), 3.15 (t, 2H, $J = 6.0$ Hz, H-2), 2.66 (d, 2H, $J = 7.2$ Hz, H-b), 2.45 (s, 3H, H-7''), 2.04 (m, 4H, H-6' and H-c), 0.82 (d, 6H, $J = 6.5$ Hz, H-d).

$^{13}\text{C NMR}$ (100 MHz, D_2O) δ 197.9 (C-a), 170.9 (C-6''), 163.9 (C-2''), 160.8 (C-4''), 144.5 (C-4'), 141.9 (C-2'), 138.0 (C-3'), 134.3 (C-5'), 114.8 (C-5''), 64.9 (d, $J = 4.2$ Hz, C-1), 50.4 (C-b), 29.5 (C-2), 26.3 (C-c), 26.2 (C-7'), 21.7 (C-d), 20.5 (C-7''), 11.2 (C-6').

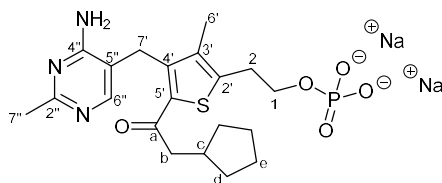
$^{31}\text{P NMR}$ (162 MHz, D_2O) δ 0.5.

HRMS (ESI) m/z : $[\text{M}+\text{H}^+]$ calculated for $\text{C}_{18}\text{H}_{26}\text{N}_3\text{O}_5\text{PS}$: 428.1409; found: 428.1418.

$^1\text{H NMR}$ in D_2O :



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(2-cyclopentylacetyl)-3-methylthiophen-2-yl}ethyl dihydrogen phosphate (5-(2-cyclopentylacetyl)-deaza-ThMP) **8f**



Prepared from **7f**. White solid (13%).

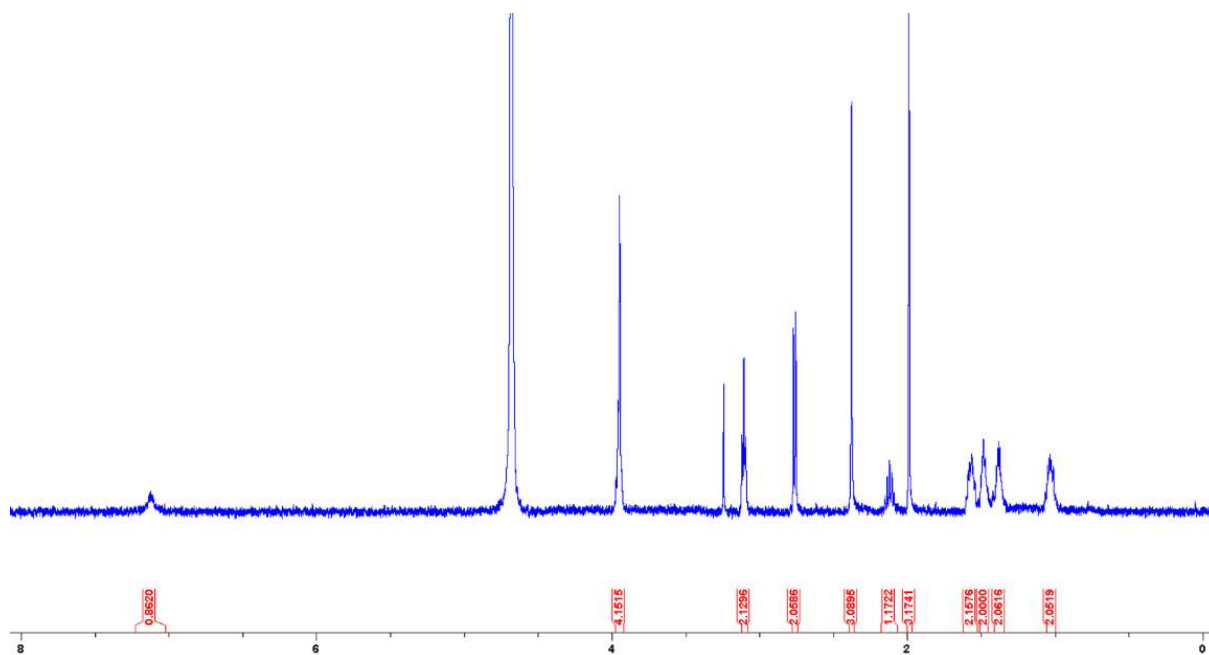
¹H NMR (400 MHz, D₂O) δ 7.13 (s, 1H, H-6''), 3.95 (m, 4H, H-1 and H-7'), 3.11 (t, 2H, *J* = 6.5 Hz, H-2), 2.76 (d, 2H, *J* = 7.4 Hz, H-b), 2.38 (s, 3H, H-7''), 2.12 (m, 1H, H-c), 1.99 (s, 3H, H-6'), 1.57 (m, 2H, H-d), 1.49 (m, 2H, H-e), 1.38 (m, 2H, H-e), 1.03 (m, 2H, H-d).

¹³C NMR (100 MHz, D₂O) δ 198.3 (C-a), 171.1 (C-6''), 163.6 (C-2''), 161.8 (C-4''), 144.6 (C-4'), 142.3 (C-2'), 137.9 (C-3'), 134.1 (C-5'), 114.6 (C-5''), 64.7 (d, *J* = 3.5 Hz, C-1), 47.7 (C-b), 37.1 (C-c), 31.8 (C-d), 29.8 (C-2), 26.2 (C-7'), 24.3 (C-e), 21.2 (C-7''), 11.2 (C-6').

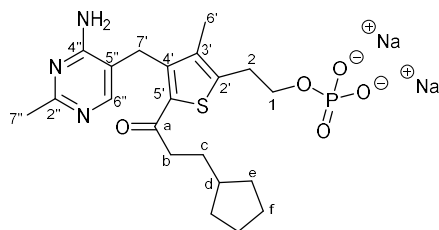
³¹P NMR (162 MHz, D₂O) δ 1.2.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₈N₃O₅PS: 454.1565; found: 454.1575.

¹H NMR in D₂O:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-5-(3-cyclopentylpropanoyl)-3-methylthiophen-2-yl}ethyl dihydrogen phosphate (5-(3-cyclopentylpropanoyl)-deaza-ThMP) **8g**



Prepared from **7g**. White solid (10%).

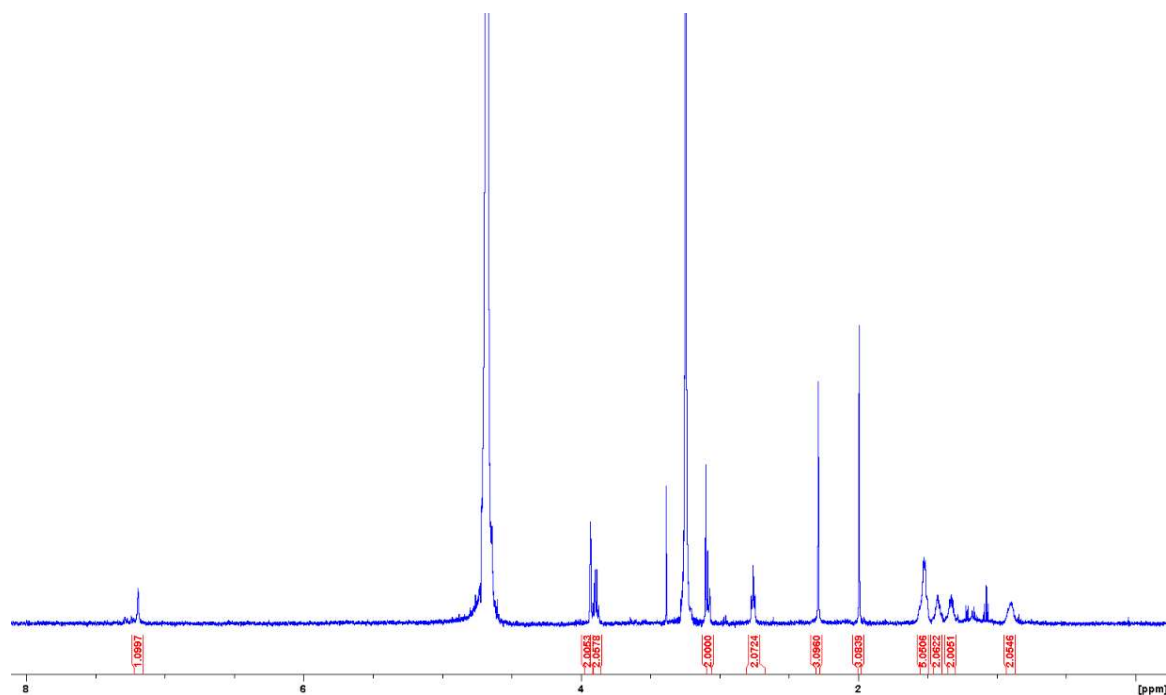
¹H NMR (400 MHz, D₂O) δ 7.20 (s, 1H, H-6''), 3.93 (s, 2H, H-7'), 3.89 (q, 2H, *J* = 6.7 Hz, H-1), 3.08 (t, 2H, *J* = 6.7 Hz, H-2), 2.76 (d, 2H, *J* = 7.3 Hz, H-b), 2.29 (s, 3H, H-7''), 1.99 (s, 3H, H-6'), 1.57 (m, 5H, 2xH-c, 1xH-d and 2xH-e), 1.43 (m, 2H, H-f), 1.34 (m, 2H, H-f), 1.03 (m, 2H, H-e).

¹³C NMR (100 MHz, D₂O) δ 198.9 (C-a), 170.6 (C-6''), 163.8 (C-2''), 162.3 (C-4''), 144.6 (C-4'), 142.9 (C-2'), 138.0 (C-3'), 133.9 (C-5'), 113.6 (C-5''), 63.7 (d, *J* = 4.2 Hz, C-1), 41.0 (C-b), 39.2 (C-d), 31.9 (C-e), 31.5 (C-2), 30.2 (C-c), 26.3 (C-7'), 24.7 (C-f), 22.8 (C-7''), 11.3 (C-6').

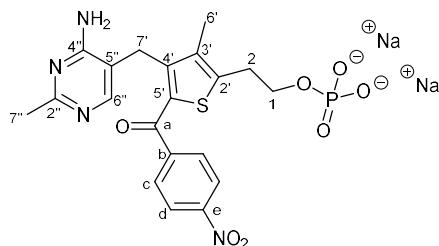
³¹P NMR (162 MHz, D₂O) δ -0.4.

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₁H₃₀N₃O₅PS: 468.1721; found: 468.1725.

¹H NMR in D₂O:



2-{4-[(4-Amino-2-methylpyrimidin-5-yl)methyl]-3-methyl-5-(4-nitrobenzoyl)thiophen-2-yl}ethyl dihydrogen phosphate (5-(4-nitrobenzoyl)-deaza-ThMP) **8h**



Prepared from **7h**. White solid (31%).

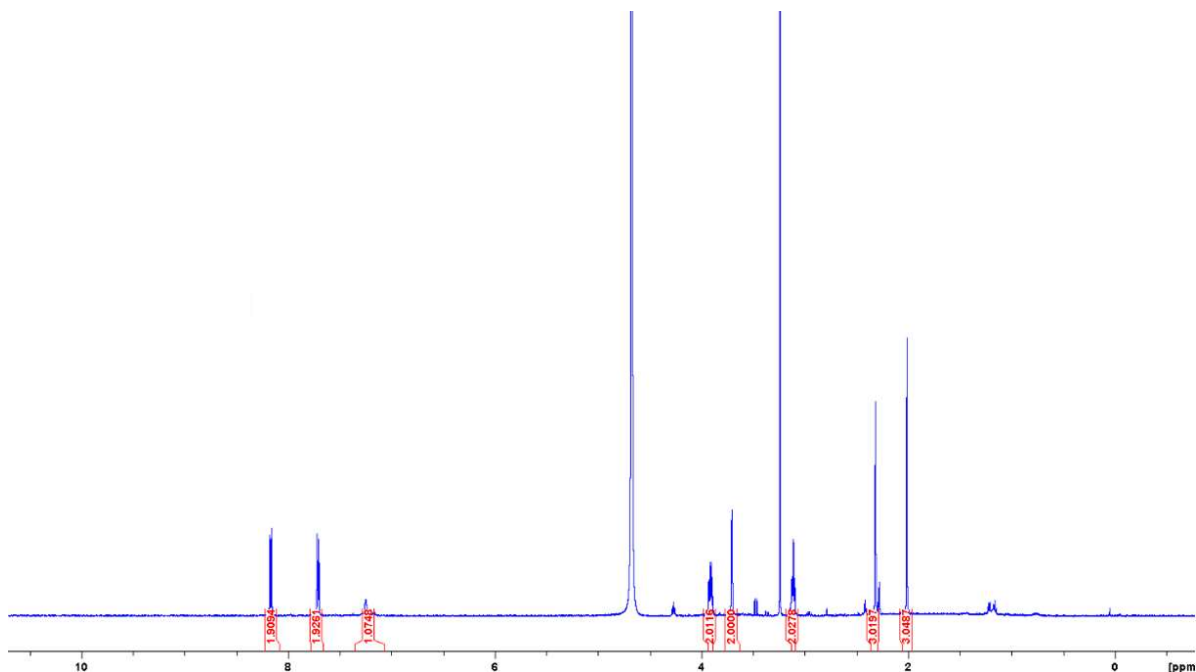
¹H NMR (400 MHz, D₂O) δ 8.17 (d, 2H, *J* = 8.7 Hz, H-d), 7.71 (d, 2H, *J* = 8.7 Hz, H-c), 7.25 (s, 1H, H-6''), 3.91 (q, 2H, *J* = 6.5 Hz, H-1), 3.71 (s, 2H, H-7'), 3.12 (t, 2H, *J* = 6.5 Hz, H-2), 2.32 (s, 3H, H-7''), 2.02 (s, 3H, H-6').

¹³C NMR (100 MHz, D₂O) δ 190.6 (C-a), 171.0 (C-6''), 162.2 (C-2''), 161.3 (C-4''), 149.4 (C-e), 148.4 (C-4'), 145.1 (c-b), 144.8 (C-2'), 138.4 (C-3'), 133.3 (C-5'), 129.2 (C-c), 123.6 (C-d), 113.4 (C-5''), 63.7 (d, *J* = 4.1 Hz, C-1), 30.1 (C-2), 26.5 (C-7'), 22.6 (C-7''), 11.4 (C-6').

³¹P NMR (162 MHz, D₂O) δ 0.15.

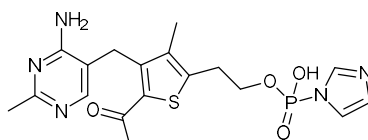
HRMS (ESI) *m/z*: [M+H⁺] calculated for C₂₀H₂₁N₄O₇PS: 493.0946; found: 493.0955.

¹H NMR in D₂O:



Preparation of 5-acetyl-deaza-ThDP 9

(2-{5-Acetyl-4-[(4-amino-2-methylpyrimidin-5-yl)methyl]-3-methylthiophen-2-yl}ethoxy)(1H-imidazol-1-yl)phosphinic acid **10**

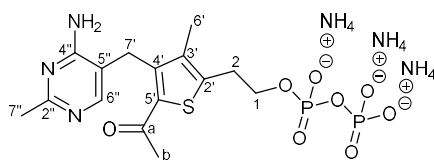


Imidazole intermediate

To a stirred solution of **8a** (86 mg, 0.2 mmol) in dry DMF (2 mL, 0.1 M) under nitrogen was added 2,2'-dithiodipyridine (88 mg, 0.4 mmol), imidazole (68 mg, 1 mmol), triphenyl phosphine (105 mg, 0.4 mmol) and TEA (0.015 mL, 0.2 mmol). The reaction mixture was stirred at r.t. for 8 h, diluted with DCM (10 mL), and extracted with water (10 mL). The aqueous phase was concentrated under reduced pressure to yield **10** as a viscous mixture, which was used in the next step without further purification.

HRMS (ESI) m/z : $[M+H^+]$ calculated for $C_{18}H_{22}N_5O_4PS$: 436.1208; found: 436.1212.

5-Acetyl-deazathiamine diphosphate (5-acetyl-deaza-ThDP) **9**



To a stirred solution of the imidazolidine intermediate **10** and zinc chloride (55 mg, 0.4 mmol) in dry DMF (1 mL, 0.2 M) under nitrogen was added tris(triethylammonium)phosphate (0.25 M in DMF, 1 mL, 0.25 mmol) dropwise. The reaction mixture was stirred at r.t. overnight and then concentrated under reduced pressure. The residue was diluted in water (1 mL) and purified by reverse-phase HPLC to yield **9** as a white solid (11 mg, 11%).

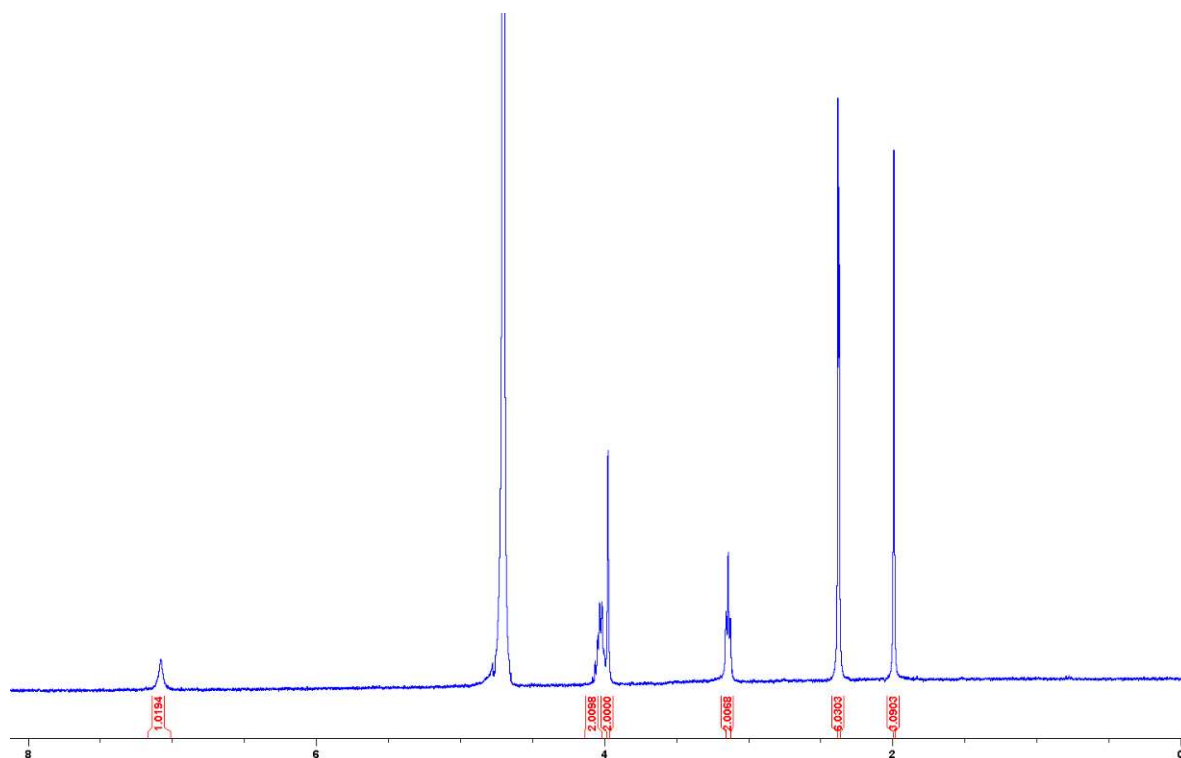
¹H NMR (400 MHz, D₂O) δ 7.08 (s, 1H, H-6''), 4.03 (m, 2H, H-1), 3.98 (s, 2H, H-7'), 3.14 (t, 2H, *J* = 6.5 Hz, H-2), 2.34-2.38 (m, 6H, H-7'' and H-b), 1.99 (s, 3H, H-6').

¹³C NMR (100 MHz, D₂O) δ 195.0 (C-a), 163.7 (C-2''), 161.1 (C-4''), 150.0 (C-6''), 144.2 (C-4'), 142.0 (C-2'), 137.7 (C-3'), 134.0 (C-5'), 114.5 (C-5''), 64.8 (C-1), 29.6 (C-2), 28.2 (C-b), 26.1 (C-7'), 20.8 (C-7''), 11.1 (C-6').

³¹P NMR (162 MHz, D₂O) δ -0.4 (d, *J* = 23 Hz), -9.7 (d, *J* = 23 Hz).

HRMS (ESI) *m/z*: [M+H⁺] calculated for C₁₅H₂₁N₃O₈P₂S: 466.0602; found: 466.0613.

¹H NMR in D₂O:



References

- 1 D. Hawksley, D. A. Griffin and F. J. Leeper, *J. Chem. Soc., Perkin Trans. 1*, 2001, 144–148.
- 2 S. Mann, C. Perez Melero, D. Hawksley and F. J. Leeper, *Org. Biomol. Chem.*, 2004, **2**, 1732.
- 3 H. Zhao, L. P. S. de Carvalho, C. Nathan and O. Ouerfelli, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 6472–6474.
- 4 L. J. Y. M. Swier, L. Monjas, A. Guskov, A. R. de Voogd, G. B. Erkens, D. J. Slotboom and A. K. H. Hirsch, *ChemBioChem*, 2015, **16**, 819–826.
- 5 A. Aponick, C.-Y. Li, J. Malinge and E. F. Marques, *Org. Lett.*, 2009, **11**, 4624–4627.
- 6 K. Ravindar, M. Sridhar Reddy and P. Deslongchamps, *Org. Lett.*, 2011, **13**, 3178–3181.
- 7 A. Iqbal, E.-H. Sahraoui and F. J. Leeper, *Beilstein J. Org. Chem.*, 2014, **10**, 2580–2585.
- 8 F. J. Leeper, A. Chan, T. Ho and D. Parle, *Org. Biomol. Chem.*, 2023, **21**, 1755–1763.
- 9 K. M. Erixon, C. L. Dabalos and F. J. Leeper, *Org. Biomol. Chem.*, 2008, **6**, 3561.
- 10 A. H. Y. Chan, T. C. S. Ho, I. Fathoni, R. Pope, K. J. Saliba, and F. J. Leeper, *ACS Med. Chem. Lett.*, 2023, in press, <http://dx.doi.org/10.1021/acsmchemlett.3c00047>
- 11 A. H. Y. Chan, T. C. S. Ho, K. Agyei-Owusu and F. J. Leeper, *Org. Biomol. Chem.*, 2022, **20**, 8855–8858.
- 12 E. P. Steyn-Parvé and H. Beinert, *J. Biol. Chem.*, 1958, **233**, 843–852.
- 13 B. Jahn, N. S. W. Jonasson, H. Hu, H. Singer, A. Pol, N. M. Good, H. J. M. O. den Camp, N. C. Martinez-Gomez and L. J. Daumann, *J. Biol. Inorg. Chem.*, 2020, **25**, 199–212.
- 14 D. A. Walsh, R. H. Cooper, R. M. Denton, B. J. Bridges and P. J. Randle, *Biochem. J.*, 1976, **157**, 41–67.
- 15 R. J. Diefenbach and R. G. Duggleby, *Biochem. J.*, 1991, **276**, 439–445.
- 16 B. Sedewitz, K. H. Schleifer and F. Götz, *J. Bacteriol.*, 1984, **160**, 273–278.
- 17 S. Takahashi, T. Kuzuyama, H. Watanabe and H. Seto, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 9879–9884.
- 18 B. Altincicek, M. Hintz, S. Sanderbrand, J. Wiesner, E. Beck and H. Jomaa, *FEMS Microbiol. Lett.*, 2000, **190**, 329–333.
- 19 R. Hamid, S. Adam, A. Lacour, L. M. Gomez and A. K. H. Hirsch, Structural analysis of 1-deoxy-D-xylulose 5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding, *bioRxiv*, preprint 2022, <https://doi.org/10.1101/2022.07.04.498669>
- 20 D. A. Clemente, A. Marzotto and G. Valle, *J. Crystallogr. Spectrosc. Res.*, 1988, **18**, 147–156.

21 D. Merk, F. Grisoni, L. Friedrich, E. Gelzinyte and G. Schneider, *J. Med. Chem.*, 2018, **61**, 5442–5447.

22 Y. Kaburagi and Y. Kishi, *Org. Lett.*, 2007, **9**, 723–726.

Chapter 2.3

Allosteric Inhibition of DXPS: A New Approach in Drug Discovery

Allosteric Inhibition of DXPS: A New Approach in Drug Discovery

Introduction:

1-Deoxy-D-xylulose 5-phosphate synthase (DXPS) is the enzyme that catalyzes the first and rate-limiting step of the MEP pathway, making it an attractive target for drug development. Efforts to target DXPS have mainly resulted in compounds that target the cofactor thiamine diphosphate (ThDP) binding site (1,2). Some of these are ThDP analogs, acting as competitive inhibitors (3–5). Fluoropyruvate has been identified as a substrate competitive inhibitor of DXPS from *Pseudomonas aeruginosa* PaDXPS (6). Additionally, another class of DXPS inhibitors, known as alkylacetylphosphonates (alkylAPs), reacts with Thiamin Diphosphate (ThDP) to form a stable phosphonolactyl (PLThDP) adduct. This process competes with the enamine intermediate formation, which occurs during the interaction of the natural substrate pyruvate with the cofactor Thiamine Diphosphate (ThDP), a mechanism that is unique to DXPS, despite being selective these specific class of inhibitors did not show anti-bacterial activity *in-vivo* (7,8). Most of other reported inhibitors of DXPS face the problem of a lack of selectivity. Therefore, there is a significant need to explore new strategies for targeting DXPS and developing a new class of inhibitors for this enzyme.

In this study, we present the co-crystal structure of DXPS with a new class of compounds that target an allosteric binding site. This discovery has the potential to revolutionize the approach to targeting DXPS.

Results and discussion:

The quest for a new class of DXPS inhibitors began with a virtual screening of our in-house library conducted by Antoine Lacour, a chemist from our team. The selection included one compound from the virtual screening, followed by the optimization of two hits. All of these compounds underwent testing against *pa*DXPS using the functional assay described in Chapter 1.

The parent hit, compound **1**, showed an IC_{50} value of 120 μ M. Interestingly, it was observed that the activity improved for compound **2**, when increasing the carbon linker length between the amide and the piperidine. This improvement was further enhanced by adding a second methylene group (Fig. 1).

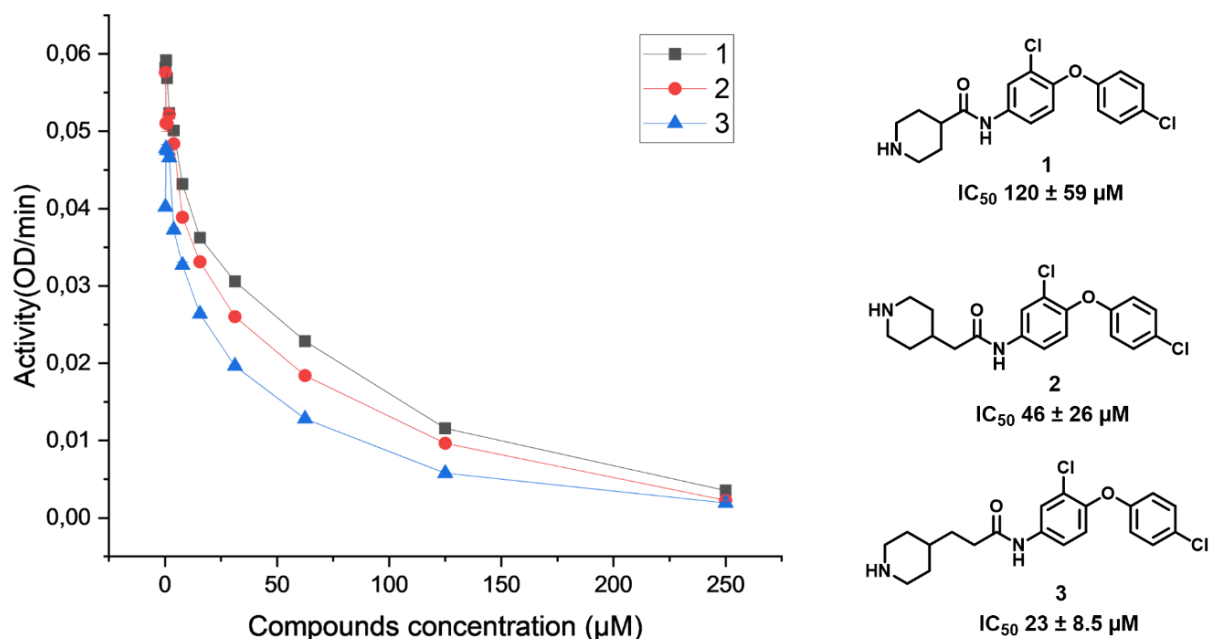


Figure 1. Chemical structures and biochemical evaluation (IC₅₀) curves of inhibitors **1**, **2** and **3**.

A mode-of-inhibition study conducted on the parent compound, compound **1**, revealed that these compounds act as competitive inhibitors with ThDP (Fig.2). We therefore anticipated binding of compounds to the ThDP binding site.

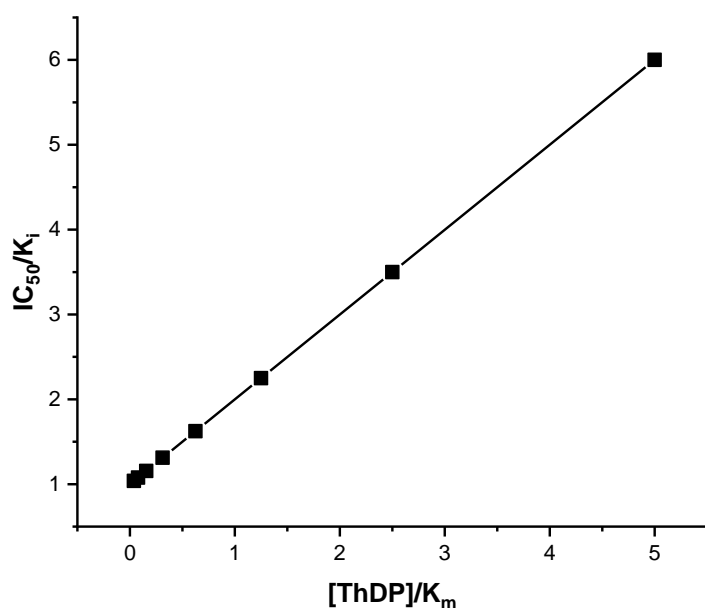


Figure 2. Exemplary IC₅₀ of inhibitor **1** with competitive modes of action as a function of substrate concentration, as predicted by the Cheng-Prusoff equation (9). A positive slope indicates increasing IC₅₀ with increasing [ThDP] suggesting compound **1** is competitive with respect to ThDP, K_m^{ThDP} 100 nM.

To further investigate the binding mode of these compounds, we decided to co-crystallize them with paXDPS using the construct, crystallization conditions, and protocol described in Chapter 1. Co-crystallization attempts were made for all compounds, but we only successfully obtained co-crystals with compound **3**, which exhibited the highest activity.

The crystals diffracted to a resolution of 1.9 Å, and the statistics of refinements are presented in Table 1. In the asymmetric unit, two units of protein were found, consistent with the expected homodimeric structure of paDXPS (10). Notably, we observed convincing electron density for compound **3** in chain A. However, chain B showed poor density for the compound, along with disordered residues surrounding the binding site.

Table 1. Data collection and refinement statistics.

Resolution range	64.39–1.91 (1.978 –1.91)
Space group	P 21 21 21
Unit cell	77.673 115.105 132.178 90 90 90
Total reflections	376956 (31862)
Unique reflections	90083 (8844)
Multiplicity	4.2 (3.6)
Completeness (%)	97.41 (97.09)
Mean I/sigma(I)	6.65 (1.31)
Wilson B-factor	22.05
R-merge	0.1675 (0.9559)
Reflections used in refinement	90073 (8844)
Reflections used for R-free	1996 (196)
R-work	0.2319 (0.3017)
R-free	0.2606 (0.3180)
CC(work)	0.941 (0.676)
CC(free)	0.926 (0.676)
Number of non-hydrogen atoms	8509
macromolecules	8483
ligands	48
solvent	0
Protein residues	1114
RMS(bonds)	0.058
RMS(angles)	1.27
Ramachandran favored (%)	96.19
Ramachandran allowed (%)	3.72
Ramachandran outliers (%)	0.09
Rotamer outliers (%)	0.46
Clashscore	5.26
Average B-factor	25.49
macromolecules	25.47
ligands	33.53

Statistics for the highest-resolution shell are shown in parentheses.

Interestingly the co-crystal showed that compound **3** binds within a crevice near a flexible loop in domain 1 rather than within the ThDP binding site (Fig. 3). Previous research from our group has indicated that this particular site of the protein is involved in conformational changes associated with cofactor binding. Specifically, the loop adopts a closed conformation induced by ThDP binding. This observation has led us to speculate that this site may play a crucial role in ThDP binding. Consequently, it presents an opportunity to interfere with ThDP binding, thereby preventing the enzyme from assuming its catalytically active conformation. This innovative approach holds the promise of selectively inhibiting DXPS, representing a novel avenue for drug development (10).

Compound **3** occupies a cleft within domain 1, which is composed of $\alpha 6$, $\alpha 7$, $\beta 2$, $\beta 3$, and $\beta 4$. The chlorine atom in the Eastern chlorophenyl moiety forms hydrogen bonds with the side chains of Gln200, Ser203, Trp250, and the main chain of Leu248. The two phenyl rings of the chlorophenyls create a hydrophobic environment, forming a semi-circular arrangement with Leu190, Phe196, Leu248, Leu199, Leu244, and Phe245.

The chlorine atom in the middle ring establishes hydrogen bonds with the main chain of Leu190 and Phe196. Notably, the ether oxygen linking the two chlorobenzene rings is solvent-exposed and does not interact with the receptor. Similarly, the amide group is predominantly solvent-exposed and occupies a relatively spacious region compared to the rest of the molecule. This area could potentially serve as a point for molecule expansion to target the deeper hydrophobic pocket beneath the carbonyl. The nitrogen atom within the piperidine ring makes contact with Glu215 and Met218. These two residues are pivotal in coordinating the phosphate moiety of ThDP in the cofactor-bound structure. They play a crucial role in the movement of the loop and the conformational changes that occur during cofactor binding. We propose that the interaction with these two residues is the primary factor contributing to the activity of this class of inhibitors (Fig.3).

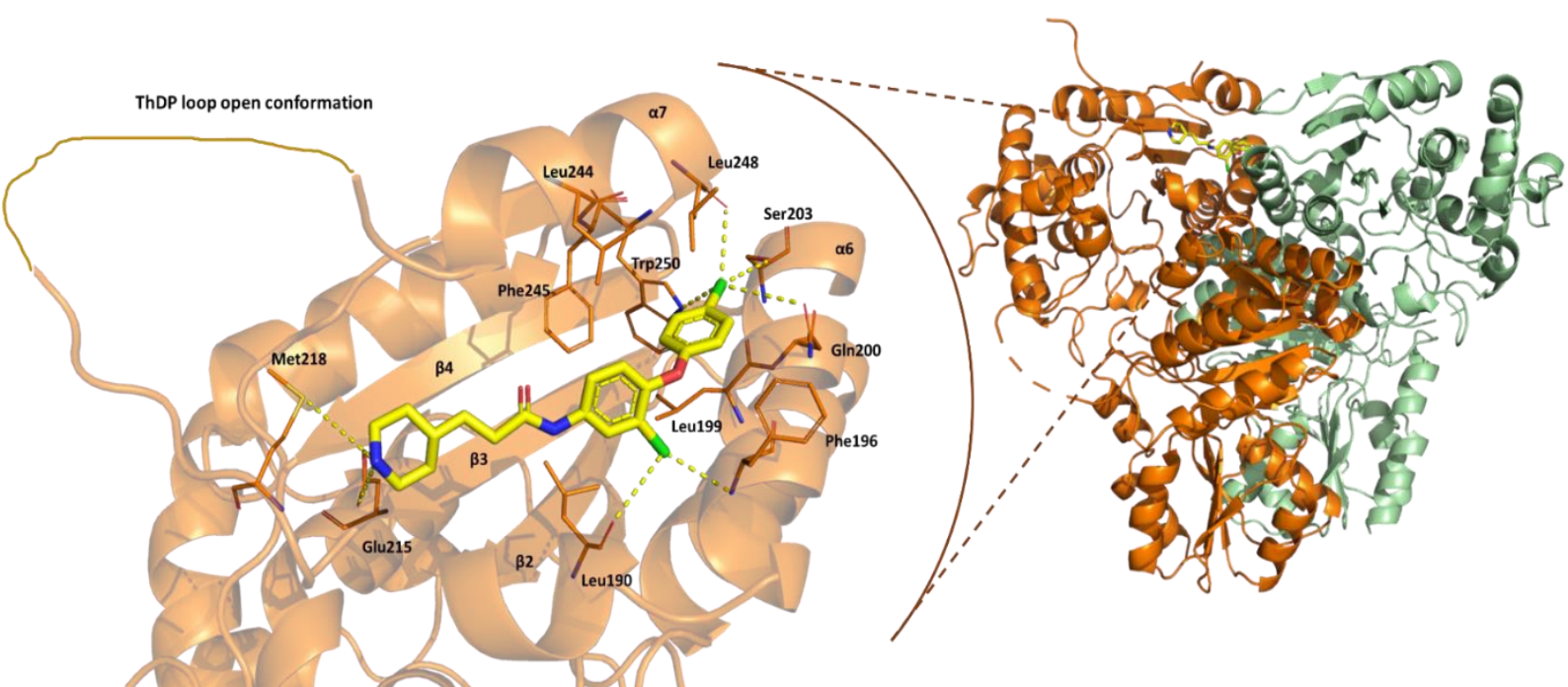


Figure 3: Co-crystal of *pa*DXPS with compound **3.** Right, dimer organization seen in the cocrystal showing chain A in orange, chain B in light green. Left, zoomed in view of the allosteric binding of compound **3**. Bonding interactions are shown in yellow dashed lines (below 5 Å). Compound and interaction residues from the receptor are shown as sticks; C, yellow and brown, O; red, N; blue, Cl, green.

Methods

Biochemical testing: The DXSP activity was analyzed at RT in a coupled assay using the protocol as described in Chapter 1. IC₅₀ values were calculated for compounds after incubating the protein with a range of concentrations between 0–250 μ M. 200 nM of paDXPS and ThDP of concentration 125 μ M. Data analysis was performed in OriginPro2023.

Crystallization: crystals were obtained after incubating the protein 15 mg/mL with 1 mM compounds in hanging drop plate, with a reservoir solution containing, 100 mM HEPES, 10% PEG-8000, and 150 mM calcium acetate. Structures were solved using molecular replacement method using the published paDXPS structure PDB ID: 8A29.

Conclusions

In conclusion, these results represent a significant breakthrough that has the potential to revolutionize the search of DXPS selective inhibitors by targeting a new allosteric binding site. This approach aims to interfere with conformational changes crucial to the catalytic event, effectively blocking ThDP binding and preventing the enzyme from assuming a catalytically active conformation.

However, it is essential to exercise caution when interpreting these data, as the observed binding in the co-crystal of compound **3** could potentially be a result of the crystallographic process. To validate the significance of this allosteric binding site and its potential as a target for drug design, further steps are necessary. Reproducing co-crystals with another compound from the same class or using the same compound with different homologs of DXPS would help confirm the validity of this allosteric pocket.

Additionally, validation experiments should be conducted, including single-point mutation studies. For instance, replacing residues like Ser203 with alanine may reduce the hydrophilicity of the site, potentially leading to the loss of critical hydrogen bonding interactions with the chlorine atom in the first chlorobenzene ring. This, in turn, could result in reduced activity of this class of inhibitors.

These findings hold promise not only for advancing our understanding of DXPS inhibition but also for opening up new avenues in the development of novel antibiotics, particularly in the fight against resistant pathogens.

References:

1. Bartee D, Freel Meyers CL. Targeting the Unique Mechanism of Bacterial 1-deoxy-D-xylulose 5-phosphate (DXP) Synthase. *Biochemistry*. 2018 Jul 24;57(29):4349–56.
2. Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate | *The Journal of Antibiotics* [Internet]. [cited 2023 Mar 9]. Available from: <https://www.nature.com/articles/ja2013105>
3. Masini T, Hirsch AKH. Development of Inhibitors of the 2C-Methyl-d-erythritol 4-Phosphate (MEP) Pathway Enzymes as Potential Anti-Infective Agents. *J Med Chem*. 2014 Dec 11;57(23):9740–63.
4. Zhu D, Johannsen S, Masini T, Simonin C, Hauptenthal J, Illarionov B, et al. Discovery of novel drug-like antitubercular hits targeting the MEP pathway enzyme DXPS by strategic application of ligand-based virtual screening. *Chem Sci*. 2022 Sep 21;13(36):10686–98.
5. Jumde RP, Guardigni M, Gierse RM, Alhayek A, Zhu D, Hamid Z, et al. Hit-optimization using target-directed dynamic combinatorial chemistry: development of inhibitors of the anti-infective target 1-deoxy-D-xylulose-5-phosphate synthase. *Chem Sci*. 2021 Jun 9;12(22):7775–85.
6. Altincicek B, Hintz M, Sanderbrand S, Wiesner J, Beck E, Jomaa H. Tools for discovery of inhibitors of the 1-deoxy-D-xylulose 5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with enzymes from the pathogenic bacterium *Pseudomonas aeruginosa*. *FEMS Microbiol Lett*. 2000 Sep;190(2):329–33.
7. Smith JM, Vierling RJ, Meyers CF. Selective inhibition of *E. coli* 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *Med Chem Commun*. 2012;3(1):65–7.
8. Bartee D, Sanders S, Phillips PD, Harrison MJ, Koppisch AT, Meyers CLF. Enamide Prodrugs of Acetyl Phosphonate Deoxy-d-xylulose-5-phosphate Synthase Inhibitors as Potent Antibacterial Agents. *ACS Infect Dis*. 2019 Jan 7;
9. Cheng Y, Prusoff WH. Relationship between the inhibition constant (K_1) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction. *Biochem Pharmacol*. 1973 Dec 1;22(23):3099–108.
10. Hamid R, Adam S, Lacour A, Monjas L, Köhnke J, Hirsch AKH. 1-deoxy-D-xylulose-5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding. *J Biol Chem*. 2023 Aug 9;299(9):105152.

Chapter 2.4

IspE Kinase Homologues: Characterization and crystal Structures enable structure-based design of novel inhibitors

IspE Kinase Homologues: Characterization and crystal structures enable structure–based design of novel anti-infectives

Rawia Hamid^{1,2}, Danica J. Walsh¹, Eleonora Diamanti¹, Diana Aguilar¹, Antoine Lacour^{1,2}, Mostafa M. Hamed¹, Anna K. H. Hirsch^{1,2*}

1. Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123 Saarbrücken, Germany.

2. Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany.

*Correspondence e-mail: anna.hirsch@helmholtz-hips.de

Submitted manuscript.

Abstract:

Enzymes of the methylerythritol phosphate (MEP) pathway are potential targets for antimicrobial drug discovery. Here we focus on 4-diphosphocytidyl-2-C-methyl-D-erythritol (IspE) kinase from the MEP pathway. We use biochemical and structural biology methods to investigate homologues from the pathogenic microorganisms; *Escherichia coli*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii*. We present the first crystal structure of the enzyme from *Klebsiella pneumoniae*. We introduce a new crystal form of IspE from *E. coli* that allows for the binding of substrate competitive inhibitors. In addition, we obtained co-crystal structures with inhibitors and performed a concise structure–activity relationship study to elucidate their binding mode. The study highlights the need for distinct inhibitor strategies due to structural differences among IspE homologues, particularly for *A. baumannii* IspE, which displays unique inhibitory profile. This study enhances our understanding of the MEP enzymes and sets the stage for structure–based drug design of selective inhibitors to combat pathogenic microorganisms.

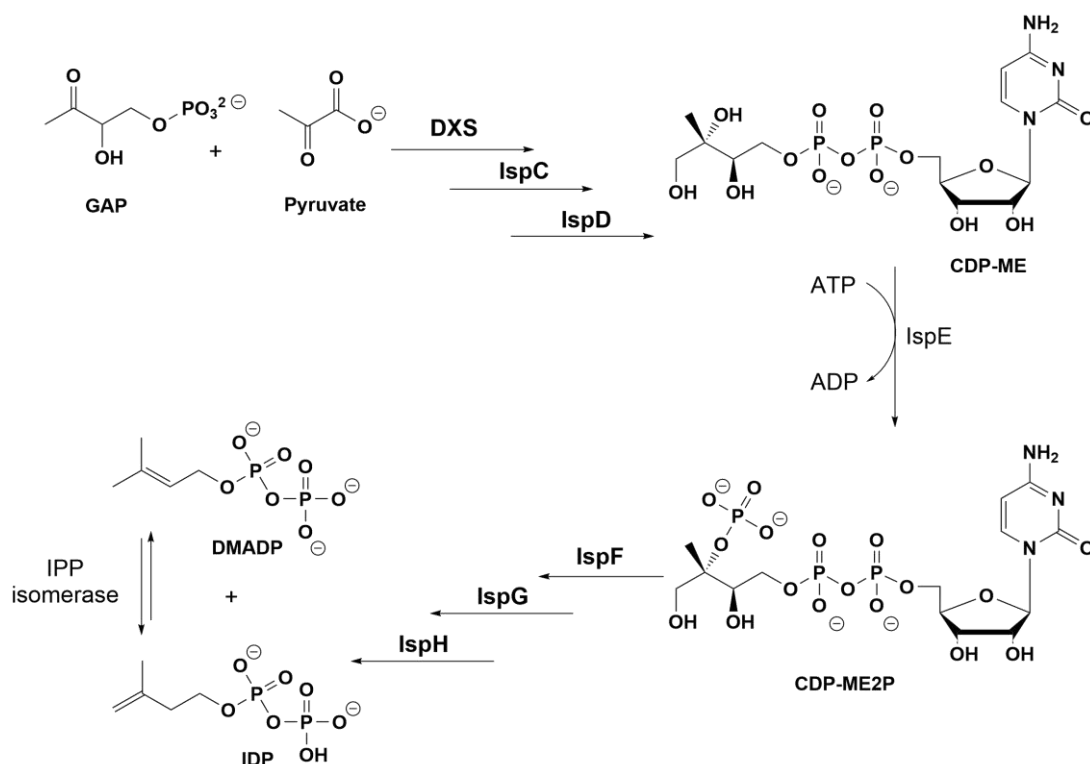
1 INTRODUCTION:

The world is racing to identify new microbial targets to combat multidrug–resistant pathogens (1, 2). The 2C-Methyl-D-erythritol 4-phosphate (MEP) pathway is a rich source of attractive underexplored anti-infective targets (3, 4) due to its role in the production of the universal precursors of isoprenoids, especially in pathogenic microorganisms (5, 6). Unlike humans, many of these microorganisms rely on the MEP pathway for the synthesis of the isoprenoids building blocks; isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP). Several Gram-negative bacteria, like *Escherichia coli*, *Klebsiella pneumoniae* and *Acinetobacter baumannii* depend on this pathway for isoprenoid production (7, 8). Constituent enzymes of the MEP pathway are validated anti-infective drug targets as confirmed through the *in-vivo* efficacy of the antimalarial drug fosmidomycin, which targets the pathway's second step catalyzed by the enzyme IspC (9, 10).

This study focuses on the only kinase present in the MEP pathway, the so-called IspE. The IspE gene (*ychB*) was first identified in 1999 by Lüttgen et al. as the enzyme that follows IspD in the MEP pathway. They provided evidence of ATP-dependent phosphorylation of the substrate 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) (11) (Scheme1). Building upon these findings, further investigations led by Kuzuyama and his team confirmed the role of the *ychB* gene and provided additional insights into its product kinetics (12). Until now, IspE crystal structures have been elucidated for *Aquifelix aeolicus* IspE (*AaIspE*, PDB ID: 2VF3) (13), *Escherichia coli* IspE (*EcIspE*, PDB ID: 10J4) (14), *Thermus thermophilus* (*TtIspE*, PDB ID: 1UEK) (15) and *Mycobacterium tuberculosis* (*MtIspE*, PDB ID: 3PYD) (16).

IspE kinase, an ATP-dependent enzyme, shares structural similarities with the galactokinase, homoserine kinase, mevalonate kinase and phosphomevalonate kinase (GHMP) enzymes. The active site of this family is characterized by five regions responsible for binding the substrate CDP-ME and co-factor ATP. These regions interact specifically with adenosine, methyl-erythritol, cytidine, and phosphate groups. Additionally, there is a small hydrophobic pocket that is partially accommodated by the methyl group of the substrate CDP-ME (13–15). Most GHMP kinase proteins exist and function as oligomeric states in solution, IspE is uniquely mostly observed as a monomer in solution (17). Nevertheless, in two available structures of *EcIspE*, potential homodimer arrangements are seen within the crystal lattice. Notably, each homodimer configuration is distinct. A structure of IspE in complex with CDP-ME and ATP (PDB-ID:10J4) has a 2-fold symmetry (14). 2WW4, a co-crystal structure of IspE and ADP, on the other hand, lacks this symmetry (18). In this structure, parts of both molecules in the asymmetric unit (ASU) are found too close to each other's active site, specifically the CDP-ME pocket, which would suggest dimer formation. However, experimental data later indicated that only a small portion of *EcIspE* forms dimers in solution (18). In addition, this proximity leads to enzyme inactivity *in vivo*. This suggests that the observed dimer formation is an outcome of the crystallization process itself. This arrangement seems to be predominant in the IspE crystal lattice, which poses a challenge in drug design. Attempting to co-crystallize/soak inhibitors targeting the substrate-binding site to elucidate their binding mode is hampered by an essentially unavailable pocket.

Here we aim to investigate the kinetic and structural properties of IspE from human pathogens, namely, *Escherichia coli* (*EcIspE*), *Klebsiella pneumoniae* (*KpIspE*), and *Acinetobacter baumannii* (*AbIspE*). Furthermore, we focus on the structural analysis of the interactions between the ligand and the active-site of IspE from these three pathogens. By investigating these aspects, the study seeks to gain a deeper understanding of the IspE enzymes in these human pathogens and their potential as targets for developing new therapeutic interventions.



Scheme1: MEP pathway, highlighting the reaction catalyzed by IspE kinase.

2 RESULTS AND DISCUSSION:

2.1 Characterization of IspE from *Escherichia coli*, *Klebsiella pneumoniae* and *Acinetobacter baumannii*

To characterize IspE kinase homologues and identify inhibitors, we first purified sufficient protein and then established biochemical assays to monitor enzyme activity. Henceforth, we obtained the plasmid containing the respective genes for expressing IspE from *E. coli* (*EclspE*), *K. pneumoniae* (*KplspE*), and *A. baumannii* (*AbIspE*) proteins from BioCat (BioCat GmbH) and transformed them into *E. coli* BL21 (DE3). Following the steps outlined in the methods section, we effectively overexpressed, purified, and confirmed proteins purity and identity. We obtained soluble, full-length IspE proteins from these pathogens. The activities of the isolated enzymes were then measured by coupling the IspE kinase activity to the auxiliary enzymes lactate dehydrogenase and pyruvate kinase, using a spectrophotometric assay that quantifies the reduction in the absorption of NADH at 340 nm (19).

EclspE $K_m^{\text{CDP-ME}}$ is 200 μM (V_{max} 171 $\mu\text{M min}^{-1}$) and K_m^{ATP} is 420 μM (V_{max} 195 $\mu\text{M min}^{-1}$), our values show the same range of kinetics values reported previously for *EclspE* (9, 12, 13). For *KplspE* the apparent $K_m^{\text{CDP-ME}}$ is 170 μM (V_{max} 119 $\mu\text{M min}^{-1}$) and K_m^{ATP} is 348 μM (V_{max} 121 $\mu\text{M min}^{-1}$). The kinetic parameter of *KplspE* are comparable to those obtained from other homologues (16, 19, 20). These values are also similar to the values obtained for its closest homologue *EclspE*. Both homologues share a sequence homology of 84%, which is also reflected in their structural similarity (discussed below). *AbIspE* shows comparable values as

well for K_m^{ATP} : 474 μM (V_{max} 153 $\mu\text{M min}^{-1}$). However, it shows slightly higher values for $K_m^{\text{CDP-ME}}$: 358 μM (V_{max} 112 $\mu\text{M min}^{-1}$) (Fig. S3). It is therefore important to note that *AbIspE* shares only 39% sequence identity with *EclspE* and 42% with *KplspE*. Fig. S1 shows alignment of the sequences from the three homologues highlighting the amino acids similarities and differences in the active site that might have contributed to the different patterns of ligand affinities between these homologues (also discussed below). Most important are the change of Pro182 in *EclspE* and *KplspE* to Glu174 in *AbIspE*, in addition to the change of Cys211 in *EclspE* and *KplspE* to Phe205 in *AbIspE*. These two residues are significant for binding to the ribose moiety and methyl, of the substrate CDP–ME, respectively.

Gel-filtration analysis shows that *EclspE* and *KplspE* both exist mainly as monomers in solution, with a small amount in an oligomeric state. *AbIspE* on the other hand exists only as a monomer in all instances (fig. S2). This observation agrees with earlier studies on other IspE kinase homologues. These studies indicated that IspE operates as a monomer, with the active site formed within a single polypeptide unit, without any evident involvement from another subunit in the catalytic process (13, 18, 21). The determined biochemical and kinetic values were subsequently used to optimize assay conditions for each homologue to evaluate inhibitory activity.

2.2 *K. pneumoniae* IspE structure shows high similarity to *E. coli* IspE

The 3D structure of *KplspE* was resolved with a resolution of 1.8 Å using molecular replacement and further refined to the highest R-work value at 19% (R-free 22%). The crystals, space group $P2_12_12_1$, and unit cell dimensions; $a=65.87$ Å, $b=76.781$ Å, and $c=98.517$ Å, with two molecules within the unit cell, and corresponds to the PDB entry: **8CKH**. Each subunit consists of two main domains that form a V-shaped cavity. This cavity accommodates the active site, which binds ATP on the N-terminal domain and the substrate CDP-ME on the C-terminal domain.

Similar to other IspE kinases with known structures, the *KplspE* structure reveals conserved sequence motifs within the GHMP kinase superfamily, specifically motifs A, B, and C. Motif A starting sheet $\beta 2$ to $\beta 3$ (Lys11 to Leu36) contains the substrate binding site, while the glycine-rich loop spanning from Gly102 to Gly108 constitutes Motif B, which interacts with the triphosphate moiety of ATP. Motif C (Leu238 to Val245) acts as a linker, stabilizing the conformation of Motifs A and B. The catalytic center results from the collective action of these motifs. (Fig. 1).

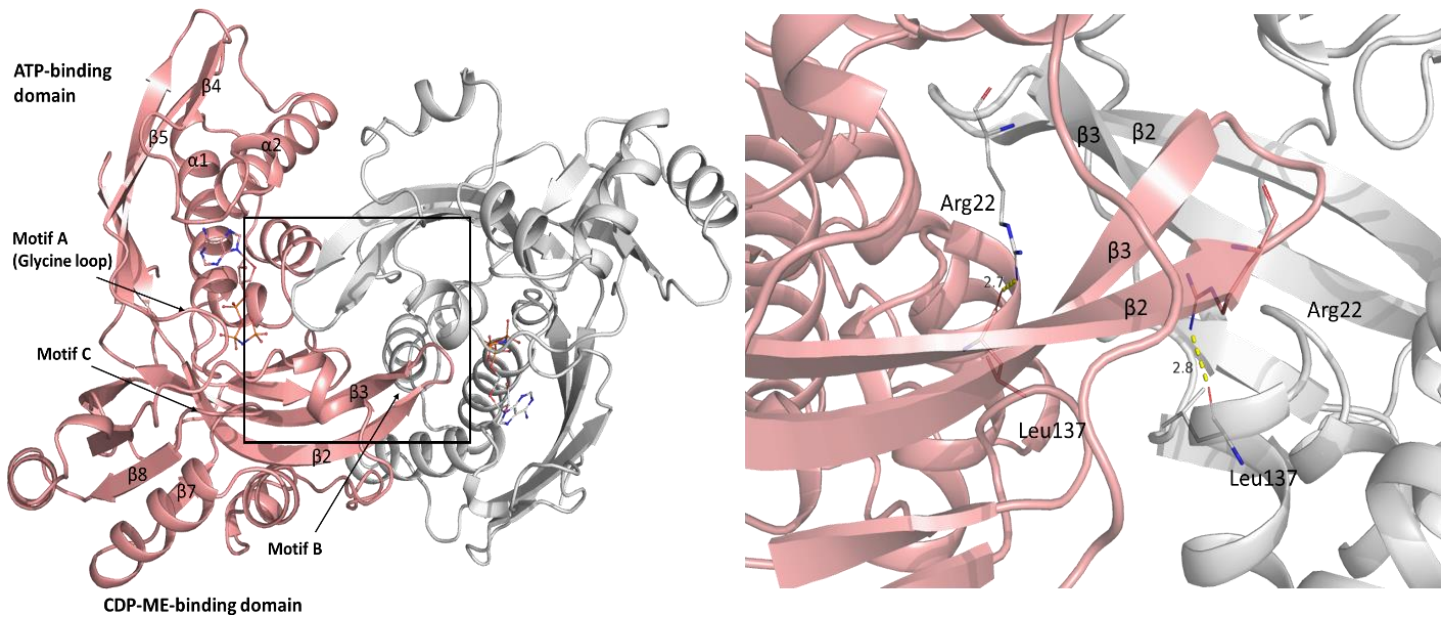


Figure 1: Structure of *KplspE*: showing subunits and domains organization, molecule A is shown in pink, molecule B in gray. **A:** domain organization: N-terminal, ATP-binding domain has four strands and three helices. The ATP binding site is bordering helices 1 and 2. The substrate-binding domain is made up of eight strands and five helices in the C-terminal domain. $\beta 2$ and $\beta 3$ in the C-terminal domain are antiparallel to each other and they are curved to form the main part of the substrate binding side. **B:** close up showing the clash the two subunits inflicted on each other where the $\beta 2$ - $\beta 3$ turn is sliding in the CDP-ME binding pocket. Hydrogen bonds are formed between Arg22 side chain of one molecule and main chain in Leu137 of the other molecule (distance 2.7 Å). Interacting residues are shown as sticks: C: pink and gray; O: red; N: blue; S: yellow.

While we could not obtain a ternary structure of *K. pneumoniae* with ATP and CDP-ME, our findings offer a satisfying model of *K. pneumoniae* IspE complexed with adenylyl-imidodiphosphate (ANP), an unhydrolyzable ATP analogue. The ATP binding site resides in a cavity within the N-terminal domain, enclosed by the loops connecting $\alpha 1$ and $\beta 5$ and $\alpha 2$ and $\beta 4$. Adenine binds within this cleft, surrounded by Val58, Val61, Ile67, Leu68, and Lys97. The ribose of ANP is solvent-accessible and directly interacts with side chain of Lys97, while the phosphates are coordinated by Gly104, Gly106, Gly108, and Ser109 (Fig. 2).

E. coli and *K. pneumoniae* both belong to the *Enterobacteriaceae* family, with an 84% sequence identity. Predictably, *KplspE* and its *E. coli* counterpart share similar structural features. Their 3D structures superimpose with an RMSD of 0.557 Å, indicating well-preserved overall folding and structural conformation (Fig. 2).

However, two notable differences between *EclspE* and *KplspE* structures stand out: the conservative replacement of Asp63 with Glu65 in the ATP binding site and Lys186 with Arg187 near the CDP-ME active site (Fig. 2D). These changes have minimal impact on the active site architecture. Additionally, in the unoccupied CDP-ME pocket of *KplspE*, Tyr25 shifts from a parallel position to the $\beta 2$ - $\beta 3$ loop to a vertical orientation relative to the same loop. Meanwhile, Phe186 adopts a perpendicular arrangement compared to its conformation in the CDP-ME-bound state (Fig. 2C). Both residues are crucial for interacting with the cytosine head, functioning as gatekeepers that facilitate site closure during substrate binding and opening during product release (21).

The 3D-structural characteristics and kinetic values of *KplspE* reveal its significant similarity to *EclspE*, suggesting these enzymes can be used interchangeably for drug-discovery projects, when prioritizing the more robust homologue is recommended. In addition, given the similarity between *E. coli* and *K. pneumoniae* counterparts, it is reasonable to assume that the binding mechanism of ligands observed in *EclspE* applies to *KplspE* as well.

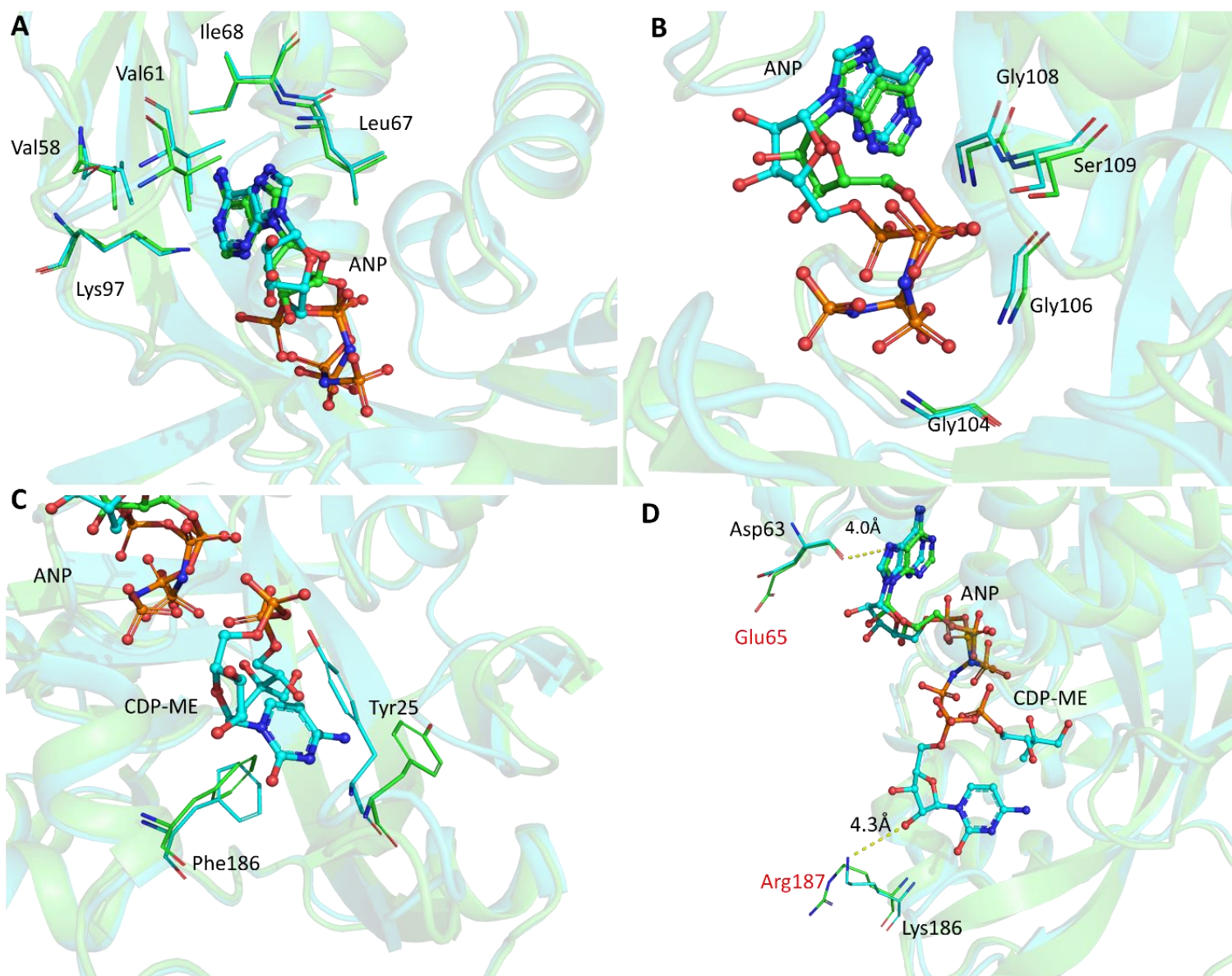


Figure 2: Comparison between *EclspE* and *KplspE*: RMSD = 0.557 (1649 to 1649 atoms) *KplspE* in green and *EclspE* cyan. **A:** Adenine binding showing the aliphatic domain and **B:** Phosphate binding domain. **C:** conformational changes in the CDP-ME binding site of Tyr25 and Phe186. **D:** view of the active site amino acids differences. Glu65 to Asp in *E.coli*, Arg187 to Lys in *E.coli*. ANP and CDP-ME atoms are shown as sticks: C: green (*KplspE*) and cyan (*EclspE*); O: red; N: blue; S: yellow.

2.3 *K. pneumoniae* IspE crystals arrangement lack symmetry and show blocked substrate binding site:

Regarding the crystal arrangement of *KpIspE*, it forms orthorhombic crystals with two molecules in the asymmetric unit (ASU). The total solvent-accessible surface area is 24,927 Å², with 1167.9 Å² (23%) as the buried interface area between the two subunits. This assembly lacks symmetry seen in the monoclinic crystals we obtained for *EclspE* (discussed below), the ASU of the orthorhombic crystal system, the secondary structure elements, specifically sheet, β 2, β 3, and the connecting turn from one molecule, slide and bind within the substrate binding pocket of another molecule. This part of the protein occupies the site to which the phosphate of CDP-ME binds, as seen in the monoclinic crystal form of *EclspE* (14). This arrangement seems to be driven by hydrogen bonds between the side chain of Arg22 in one molecule and the main chain of Leu137 in the other molecule, with an average distance of 2.7 Å (Fig. 1).

As previously mentioned, *KpIspE* primarily exists as a monomer in solution, and this assembly would render the enzyme inactive under natural conditions due to the unavailable substrate pocket. This suggests that this crystal formation occurs during the crystallization process and is not a genuine dimer. Similar crystallization artifacts have been observed before in the triclinic form of *EclspE* (18). This observation holds particular significance in drug design, especially when attempting to co-crystallize inhibitors targeting the substrate-binding pocket.

2.4 Monoclinic structure of *E.coli* IspE: a novel crystal form exhibiting a distinct arrangement of the two subunits that allows favorable ligand binding

Co-crystallization of IspE inhibitors has proven to be challenging, especially for IspE homologues that have a conformation preventing the binding of substrates or ligands designed for the substrate pocket (22). In case of *EclspE* and *KpIspE*, most of the obtained conditions resulted in crystals lacking symmetry between the two subunits, likely due to crystallization artifacts. These artifacts caused clashes that obstructed the substrate-binding pocket (18).

Essentially, the binding sites of these two subunits were positioned in opposite directions, undergoing a 180° rotation from each other. This configuration created a fully accessible pocket for ligand binding, whether through soaking or co-crystallization (Fig.3). In the new monoclinic crystal form, space group C_{121} of *EclspE* is seen, the crystals diffracted to a resolution up to 1.5Å (PDB: **8QC7**, **8QCC**, **8QCN** and **8QCO**), refinement statistics can be found in Table S2. The asymmetric unit consisted of two molecules, labeled A and B. The obtained structure aligns with the overall structure identified for *EclspE*, showcasing the distinct fold of the GHMP family (23). A similar domain arrangement to *EclspE* is also evident. The CDP-ME binding site comprises an α -helix and four anti-parallel β -sheets on one side, with four helices on the opposite side. The ATP-binding site is made up of four antiparallel β -sheets and five helices (14, 18). Although we used the unhydrolyzable ATP analog, ANP, to form the crystals, clear density was observed for only two phosphates and not the terminal phosphate, most probably because of the flexibility of the triphosphate. Consequently, we placed ADP in the

pocket instead of ANP. The binding of both molecules closely resembles that described previously for *EclspE* (14).

The total solvent-accessible area of the newly obtained crystals is 25,127 Å² and the buried interface area is 97.0 Å² (3%), compared to 1167.9 Å² (23%) in the triclinic structure and 568.7 Å² (12%) in the monoclinic $P 1 2_1 1$. The complex formation significance score (CSS) obtained from PisaServer analysis for the new conformation of *EclspE* was 0.0 (0.19 for the *KplspE* orthorhombic crystal). In addition to the gel-filtration analysis results, the achieved CSS provides more proof that the interface does not play any role in complex formation and seems to be a result of crystal packing only. (24). Fig. 3 shows the different crystal forms identified for *EclspE* in comparison to those obtained in this study.

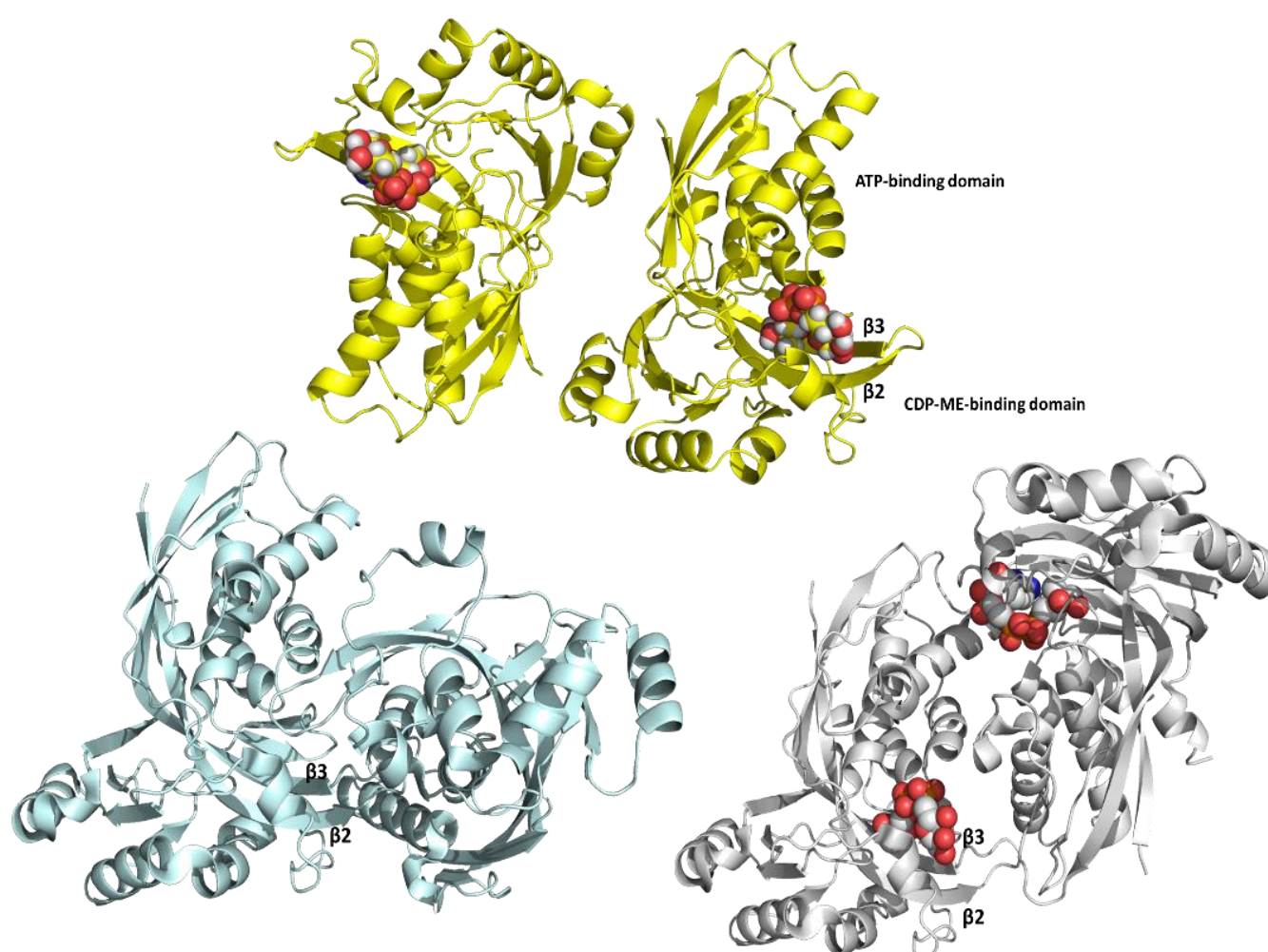


Figure 3: Comparison of the different crystal conformations of *EclspE* in the ASU: shows the difference between the two monoclinic crystal forms. PDB: **8QC7** (monoclinic C_{121}) in yellow. *EclspE*-ANP and CDPME, PDB:1OJ4 (monoclinic, $P 1 2_1 1$) in gray and the triclinic $P1$, (PDB: 2WW4) light blue, molecule A in all crystals is shown in the same orientation to show the substrate CDP-ME availability/blockage. CDP-ME is shown as spheres.

2.5 Biochemical evaluation of IspE substrate-competitive inhibitors:

In the search for inhibitors against IspE, it is possible either to target the ATP or the substrate-binding sites. Kinase inhibitors are mostly designed to target the ATP-binding site (25, 26). In IspE kinase however, this pocket is rather shallow and mostly solvent-exposed. Over the years, our group attempted to address this pocket. We, recently, performed a virtual screening on the ATP binding pocket followed by an SAR study (27), and in the past Tang et al. used a library of existing GHMP kinases inhibitors (28). Unfortunately, both attempts were not very successful in developing selective IspE inhibitors targeting the kinase ATP pocket.

The substrate-binding pocket and subpockets on the other hand have been successfully targeted (22, 29, 30). The work published in 2007 led to the identification of compound (\pm)-1 (29, 31). Brenk group also extensively focused on the study of IspE inhibitors and they reported an *in silico* and HTS approaches that identified inhibitors in the high micromolar range (32). Therefore, it is evident that more efforts are needed towards the optimization of the existing inhibitors besides the identification of novel chemical classes targeting the enzyme IspE. To address this challenge further insights about the binding mode are required and for this we used here compound (\pm)-1 (Fig. 4) as a reference inhibitor due to its high potency and selectivity against IspE. As shown in Fig. 4, this structure consists of a central cytosine scaffold (A) bearing a tetrahydrothiophenyl ring (B) as ribose analogue bridged via a triple bond (C) to a cyclopropyl sulfonamide (D). Aiming to identify determinants for IspE inhibition by (\pm)-1, we conducted a focused SAR study that separately targeted each component of the molecule.

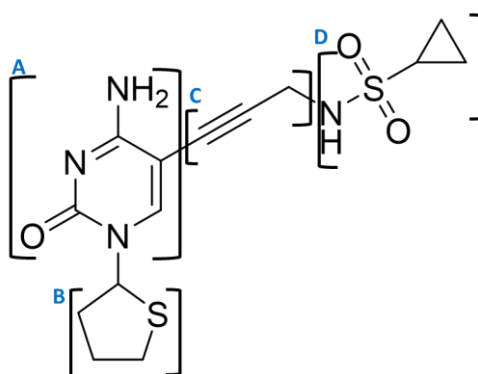


Figure 4. Chemical structure of compound (\pm)-1, illustrating the SAR strategy.

The compounds were simultaneously tested *in vitro* on the three IspE homologues studied in this work (Fig. S4). The homologues exhibited an inhibition pattern that correlates with their structural similarities and differences in the binding site, indicating the need for different strategies in designing inhibitors targeting each of these enzyme homologues. Compounds (\pm)-1–6 showed similar inhibition values for *Ec*IspE and *Kp*IspE. While compounds (\pm)-1–4 retained activity against those, only compounds (\pm)-1 and 2 retained activity against *Ab*IspE. The designed and synthesized subset of compounds (1–6) and their biological evaluation are depicted in Fig. 5.

Next, we selected the most potent and soluble compounds for co-crystallization to further confirm their binding mode and guide our optimization strategies. Although we could get high-resolution co-crystal structures with *EclspE*, the crystal conformation hindered the possibility of obtaining a co-crystal with *KplspE*. Nevertheless, due to the high similarity observed between the homologues from *E. coli* and *K. pneumoniae*, we propose the binding mode can be translated from *EclspE* to *KplspE*, which is also supported by the comparable IC_{50} values of the tested inhibitors. On the other hand, all our efforts have failed in obtaining a co-crystal structure with *AbIspE*, due to its high solubility, which means that high concentrations are required to drive crystallization. In addition, because *AbIspE* has 13% lysine residues in its sequence, mostly on the surface, we tried lysine methylation which resulted in poor-quality crystals. However, we used sequence alignments and a homology model of *AbIspE* to try to rationalize the differences in the inhibitory activity reported here.

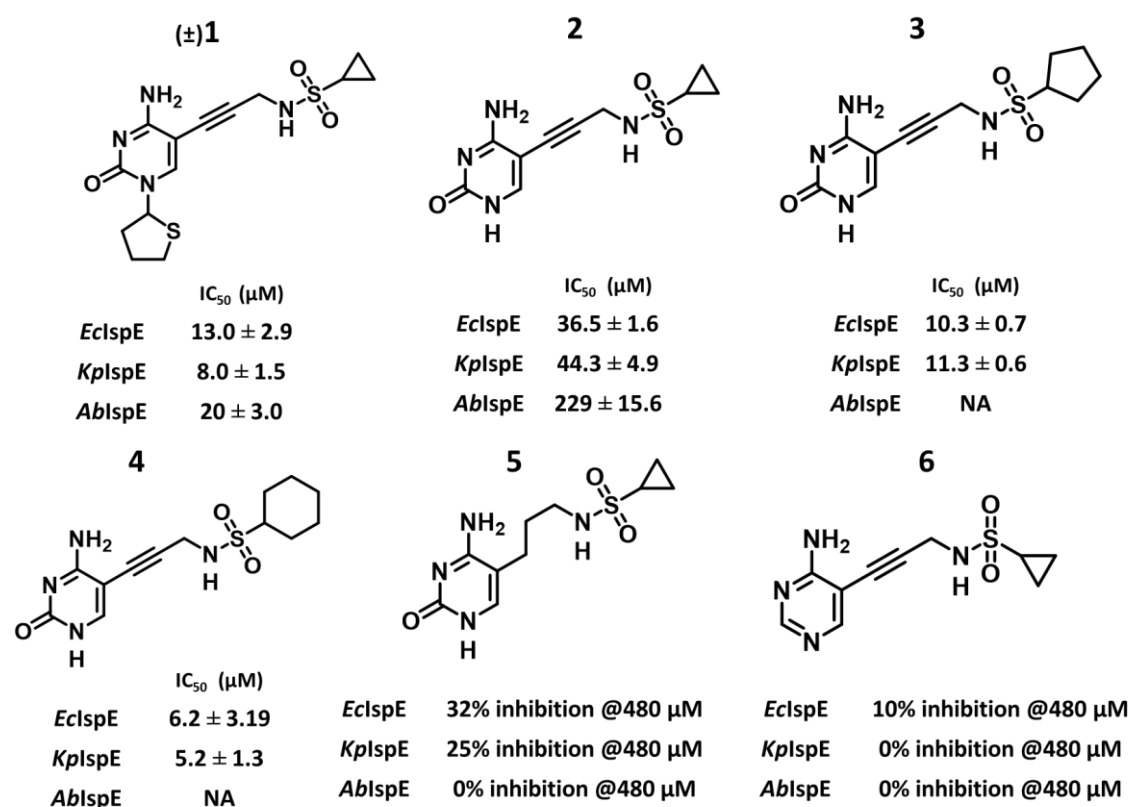


Figure 5: Biochemical evaluation of compound (\pm)1 derivatives. The half-maximal inhibitory concentration (IC_{50}) on *EclspE*, *KplspE* and *AbIspE*. The values presented are average of duplicate measurements. NA: not available; no dose-response curve was obtained to calculate the IC_{50} value.

2.6 Co-crystallization of IspE inhibitors:

In the second part of our study, we focused on investigating the binding modes of these compounds by co-crystallizing them with *EclspE*. It is important to note that we achieved the novel conformation of *EclspE* crystals that is amenable to soaking only after incubating the protein with a significantly high amount of the substrate. This approach aimed to separate the two subunits and prevent them from adopting the clashed conformation, thereby freeing the

CDP-ME pocket. Consequently, these crystals are suitable only for soaking with compounds capable of displacing the bound CDP-ME. Furthermore, these compounds must be soluble enough to be used at notably high concentrations for this purpose. The substrates of the MEP pathway, especially IspE, exhibit high polarity due to their di- and triphosphate groups (33). As a result, the active site is highly polar, which in turn presents a challenge in designing potent inhibitors. However, in our specific context, this polarity translates to high solubility for the selected compounds, allowing us to use the required high concentrations of these compounds for soaking the CDP-ME-bound crystals.

As mentioned earlier, our choice was directed towards (**±**)**1** and its derivatives. In all the structures obtained with *E*lspE and compounds **1** (PDB: 8QCC), **2** (PDB: 8QCN), and **3** (PDB: 8QCO), a consistent interaction pattern was observed, similar to the interaction observed with the cytosine ring in CDP-ME. Specifically, the central cytosine in the inhibitors is positioned between Tyr25 and Phe185, leading to the formation of hydrogen bonds with both the backbone and side chain of His26 (Fig. 6). In the co-crystal with compound **1**, the tetrahydrothiophene ring overlaps with the ribose moiety of CDP-ME, where it occupies a central position between Tyr25 and Pro182, forming a pseudo- π sandwich (Fig. 6A). While these residues are conserved in the *K*lspE structure, Pro182 is replaced by a Gln174 in *A*lspE, which means a different type of interaction will have to take place in this position. Additionally, the sulfur atom of the tetrahydrothiophene ring could potentially point to form hydrogen bonds with Thr181 or Tyr25. The alkyne linker guides the sulfonamide moiety into the notably polar phosphate binding region situated between the substrate and the ATP binding sites. The nitrogen atom of the sulfonamide forms a hydrogen bond with the side chain of Asp141. Furthermore, the sulfone group, in relation to the NH group, adopts a staggered conformation, facilitating hydrogen bond interactions with the side chain of Asn12 and Lys10. Meanwhile, in compounds **1**, **2** and **3**, the orientation of the cyclopropyl and cyclopentyl rings is directed towards a small hydrophobic pocket delineated by Leu15, Leu28, and Phe32. This hydrophobic region partially contributes to binding the methyl group of the substrate, CDP-ME. Our experimental findings confirm these compounds are indeed binding to the substrate-binding pocket as intended, forming interactions similar to those observed with the native substrate (Fig. 6A&7).

Compound **2** was designed to experimentally evaluate the importance of the tetrahydrothiophene ring. Interestingly, it demonstrated an almost three-fold decrease in activity with IC₅₀ values of 36.5 μ M and 44.3 μ M against *E*lspE and *K*lspE respectively, and a 10-fold decrease with an IC₅₀ of 229 μ M against *A*lspE. The co-crystal structure shows that this drop in activity was due to the loss of the tetrahydrothiophene ring's hydrophobic interaction (stacked between Tyr25 and Pro182) in addition to losing the potential hydrogen bond of the S atom with Thr181 or Tyr25 (Fig. 6B). Despite the dramatic decrease in activity especially against *A*lspE, the compound still retained some activity against the other homologues. Indicating that the rest of the molecule is still contributing significantly for binding.

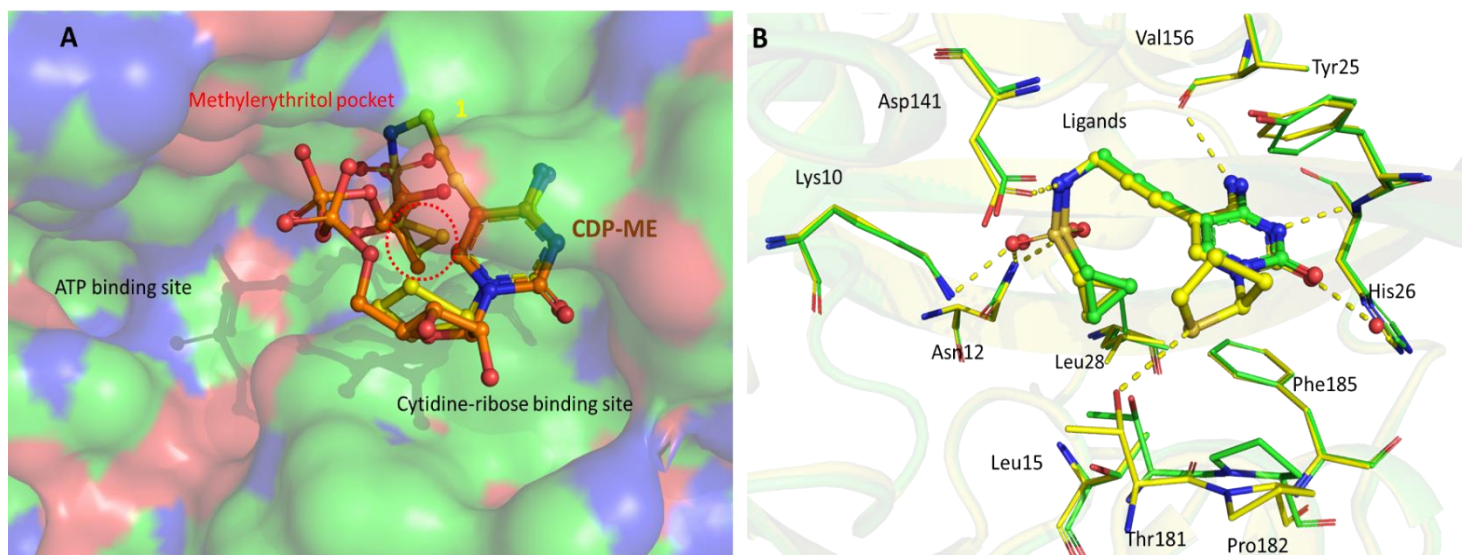


Figure 6: A: Surface representation of *EclspE* active site, showing an overlay between the CDP-ME and the competitive inhibitor (\pm)1 in the substrate pocket. IspE Active site is composed of three main pockets: the cytidine binding pocket with the ribose sub-pocket, the methylerythritol (ME) binding site featuring a small hydrophobic cavity (highlighted by a red circle), and the ATP binding pocket. **B: Superimposition of experimental structures of *EclspE* co-crystallized with compounds (\pm)1 in yellow and 2 in green showing binding interactions discussed in the text.** Molecular surface representations performed using PyMol. Color code: C: green and yellow; O: red; N: blue; S: yellow.

We designed and synthesized compounds **3** and **4** with bigger ring size in place of the cyclopropyl to test the optimum binding occupancy of this small hydrophobic pocket. Both compounds were tested on all homologues and co-crystallized with *EclspE*. The results showed that both compounds have regained the lost activity in compound **2** and slightly improved the affinity compared to (\pm)**1**. Compound **4** with the six-membered cyclohexyl ring additionally improved the activity further than compound **3** with the five-membered cyclopentyl ring. Theoretical and experimental studies have previously shown that occupation of hydrophobic pockets in a protein target is the main contributor to the binding affinity of the ligand (34). In IspE, it was found that this hydrophobic pocket, originally accommodating the methyl group of the methylerythritol moiety, is notably more spacious than initially anticipated (Fig. 7). Therefore, the appropriate filling of this sub-pocket would result in a significant gain in binding free enthalpy (35). The volume of this pocket measures roughly around 180 \AA^3 in *EclspE* and *KplspE* and 102 \AA^3 in *AbIspE*. Prior studies indicated a significant relationship between its occupancy and activity of the ligands targeting the substrate pocket (30). Although we could not obtain a co-crystal structure with compound **4**, due to solubility issues, a co-crystal structure of compound **3** with *EclspE* was obtained, which showed that the cyclopentyl ring indeed occupies the deep hydrophobic pocket surrounded by nonpolar residues, Ile15, Ile17, Leu28 and Ph185 (Fig. 7). This hydrophobic interaction exhibited sufficient strength for these two compounds to possess higher affinities than compound (\pm)**1**. This factor should be taken into account when considering an extension from this sub-pocket to the adjacent phosphate-binding pocket.

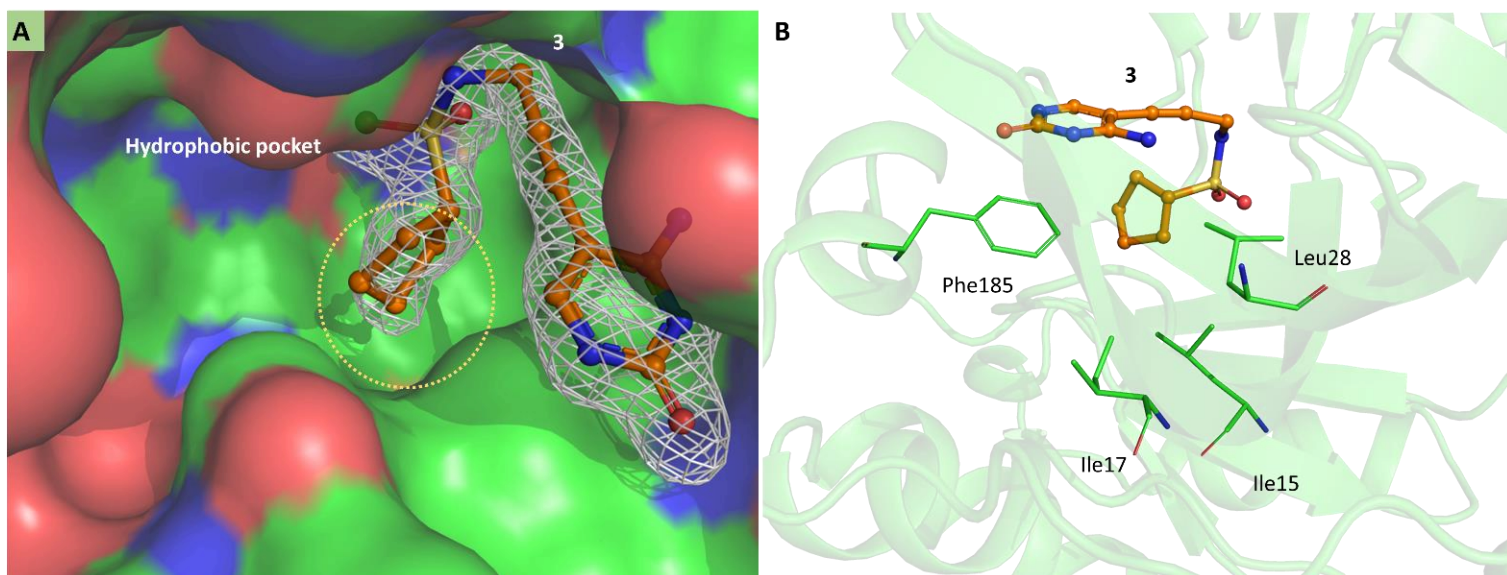


Figure 7: Binding of ligand 3 to EclspE (PDB: 8QCO): **A:** Molecular surface representation of EclspE showing the hydrophobic pocket occupied by the five-membered hydrocarbon ring from the ligand. **B:** Close-up of the small, hydrophobic pocket showing residues of the receptor and the ligand. Residues are shown as sticks. Color code: C: green; O: red; N: blue; S: yellow.

2.7 *A. baumannii* IspE requires distinct targeting strategies compared to *E. coli* and *K. pneumoniae* IspE.:

AbIspE appears to show inhibition pattern comparable to the other tested homologues relative to compound **1**. The inhibition values for **2**, **3** and **4** however, seem to be substantially different on *AbIspE*. We could not obtain X-ray crystal structure with *AbIspE* but we used sequence alignment and homology modeling to try to justify this and to structurally showcase the differences.

AbIspE (Uniprot accession code : B2HUM5). The Swiss-Model structure was downloaded as a PDB file from the Swiss-Model website via the link provided on the Uniprot web page (36). The structure was then aligned with the *EclspE* structure in MOE with an overall RMSD of 1.118 Å (37). Fig. 8, shows the structural differences between the two homologues within the CDP-ME binding site.

Compounds **2**, **3**, and **4** were then docked into the CDP-ME binding site of interest of *EclspE* (experimental crystal structure) and *AbIspE* (SwissModel structure). We were interested in rationalizing the observed differences in activity arising from the different ring sizes at the sulfonamide moiety, the docked poses of **2**, **3** and **4** closely match the crystallographic binding mode in the experimental *EclspE* structures. The docking scores of the three compounds are shown in Table S2 Gratifyingly, the docking scores closely match the experimental activities, with **2** showing a slightly lower predicted binding energy than **3** and **4**. We attribute this to the reduced ability of the cyclopropyl ring to interact with the small hydrophobic pocket at the back of the binding site. The cyclopentyl and cyclohexyl rings seem to fit this small

hydrophobic pocket better and this is reflected in both the docking scores and experimental potencies.

On the other hand, when looking at Fig. 9A, B, and C, we observe significant differences of the binding mode of these compounds in the *AbIspE* structure. For **2** (Fig. 9A), there is loss of the interaction of the cytosine carbonyl with His25 as well as the loss of the interaction between the sulfonamide oxygen and the side chain of Asn11 (when compared to the *EclspE* binding mode). Similarly, for **3** (Fig. 9B), we observe the loss of hydrogen bonding interactions between the cytosine ring and the backbone of Ile48 with the formation of an internal hydrogen bond instead. Finally, for **4** (Fig. 9C), the interactions of the cytosine ring with the backbone are maintained, however, in the docked pose, the sulfonamide no longer interacts with the side chains of Asn11 and Asp133 leading to a loss in predicted binding energy relative to the *EclspE* structure. From these results, we conclude that the replacement of Cys211 in *EclspE* with Phe205 in *AbIspE* reduces the size of the small hydrophobic pocket accommodating the cyclopropyl ring of **3** (Fig. 9A). This causes a shift in the binding mode of these compounds in the pocket of *AbIspE*, leading to the loss of important hydrogen bonding interactions made by either the cytosine ring or the sulfonamide moiety. This explains the observed drop in experimental activity of this series of compounds on the *AbIspE*.

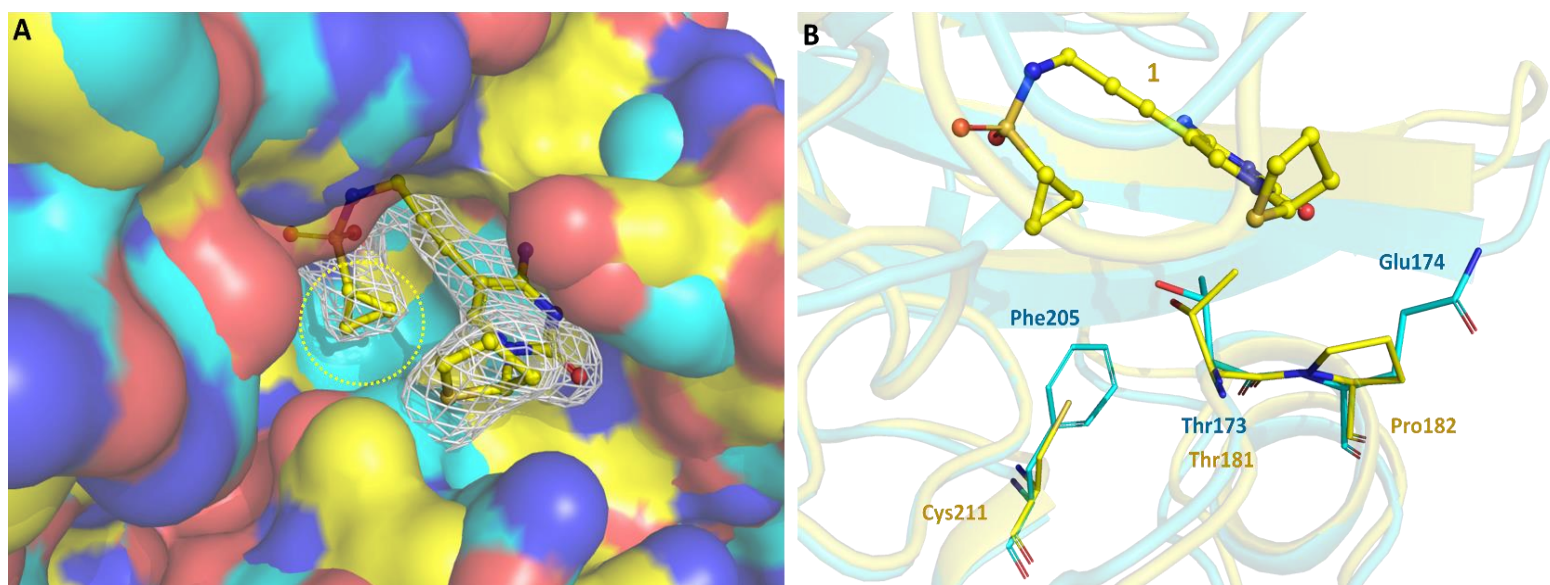


Figure 8 : *EclspE* binding to **1 superimposed with homology modeled structure of *AbIspE*.** *EclspE* in yellow superimposed with *AbIspE* in cyan. **A:** molecular surface representation of the *AbIspE* showing the tighter hydrophobic pocket targeted by the cyclopropane in the ligand. **B:** Showing a close-up of the main difference in the active site between the two homologues. In which a cysteine (Cys211) in *EclspE* is replaced by a phenylalanine (Phe205) in *AbIspE*. Color code: C: cyan and yellow; O: in red; N: in blue; S: yellow.

Few additional findings are worth mentioning; We tested the importance of the alkyne linker by synthesizing compound **5** with a flexible linker. This resulted in a complete loss of activity across all homologues suggesting the important role of this alkyne linker to direct the cycloalkyl rings into the hydrophobic pocket. The study also highlighted the challenge of cytosine replacement without compromising ligand affinity, a contrast to the numerous adenine substitutions seen in kinase inhibitors. The replacement of cytosine with just an amino pyrimidine ring resulted in a significant reduction in activity as observed in compound **6**. These preliminary findings will guide our future synthesis in the development of novel IspE inhibitors.

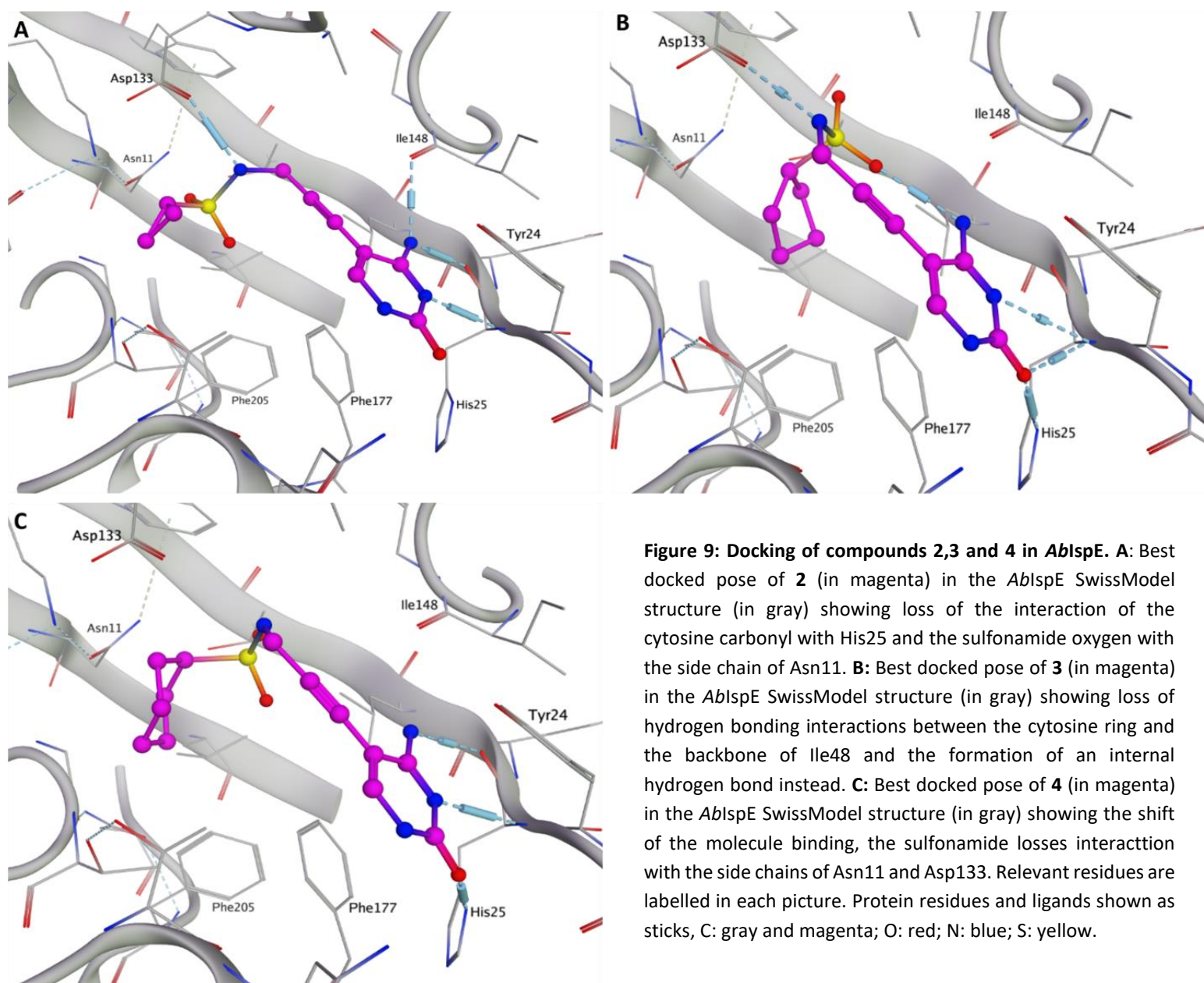


Figure 9: Docking of compounds 2,3 and 4 in AblSpE. **A:** Best docked pose of **2** (in magenta) in the AblSpE SwissModel structure (in gray) showing loss of the interaction of the cytosine carbonyl with His25 and the sulfonamide oxygen with the side chain of Asn11. **B:** Best docked pose of **3** (in magenta) in the AblSpE SwissModel structure (in gray) showing loss of hydrogen bonding interactions between the cytosine ring and the backbone of Ile48 and the formation of an internal hydrogen bond instead. **C:** Best docked pose of **4** (in magenta) in the AblSpE SwissModel structure (in gray) showing the shift of the molecule binding, the sulfonamide loses interaction with the side chains of Asn11 and Asp133. Relevant residues are labelled in each picture. Protein residues and ligands shown as sticks, C: gray and magenta; O: red; N: blue; S: yellow.

3 CONCLUSIONS

Our study aimed to characterize IspE kinase homologues, explore the substrate pocket and study inhibitor binding modes. To achieve this, we recombinantly overexpressed and purified the proteins from *EclspE*, *KplspE*, and *AbIspE*, established biochemical assays, and obtained 3 co-crystal structures of *EclspE* with inhibitors.

EclspE exhibited kinetics consistent with previous reports, while *KplspE* parameters aligned with those of similar homologues. *AbIspE* kinetic values highlighted its divergence resulting a lower sequence identity. These parameters were central in optimizing the assay conditions for each homolog, setting the stage for high-throughput screening in search for inhibitors.

We presented the first crystal structures of *K. pneumoniae* IspE in complex with the non-hydrolyzable ATP analogue, providing valuable insights into its interactions. Despite the distinctive arrangement and interactions in the crystal lattice of *EclspE* and *KplspE*, gel-filtration analysis showed that *KplspE* exists mainly as a monomer in solution. The 3D structural characteristics and kinetic parameters of *KplspE* closely resemble those of *EclspE*. Therefore, we suggest that they can be employed interchangeably, particularly when encountering challenges in obtaining and/or crystallizing either enzyme. This proposition relevant in light of our discovery of a new crystalline form for *EclspE*, which allows ligand binding during soaking—a process previously proven to be difficult. We demonstrated the applicability of this condition by successfully obtaining three new co-crystal structures of *EclspE* with inhibitors.

We observed that *AbIspE* necessitates distinct inhibitor design due to structural dissimilarity, with a smaller hydrophobic subpocket influencing compound affinities. While the strategy for targeting the substrate-binding site in *EclspE* and *KplspE* might involve proper filling of a hydrophobic pocket and covalent targeting of a cysteine residue deep in the CDP-ME pocket. For *AbIspE*, the attention could be drawn to the other side of the pocket where a glutamine is replacing a proline, that is important for interaction of *EclspE* and *KplspE* with the ribose moiety in CDP-ME or ribose substituents in ligands designed to target the substrate pocket.

Our findings emphasize the nuanced nature of IspE kinase inhibition, guided by subtle structural variations among homologues and offer valuable insights into the development of potent and selective inhibitors for IspE kinase homologues, contributing to the broader goal of combatting pathogenic microorganisms reliant on the MEP pathway.

4 MATERIALS AND METHODS

Protein preparation: We transformed pET28a plasmids containing our constructs into electrocompetent BL21 (DE3) *E. coli* cells. The transformed cells were subsequently cultured on LB-Agar supplemented with 25 µg/mL kanamycin and allowed to incubate overnight at a temperature of 37° C. We then transferred the selected colonies into LB medium

supplemented with 25 µg/mL kanamycin, at 37° C until OD₆₀₀ 0.3–0.6 for *EclspE* and *KplspE*, and 0.8 for *AbispE*. We induced overexpression with 1 mM IPTG/18° C/180 RPM/overnight. The harvested cells were then disrupted before subsequently purifying the proteins. First, affinity chromatography using HisTrap HP 5 mL, then removing tags with TEV protease (1:20). The output from the reverse IMAC column was injected into a 7 mL loop of the ÄKTA pure system on an S200 SEC column. Peak fractions were combined and concentrated in the respective storage buffer for each protein. Finally, we performed SDS-PAGE gel electrophoresis to confirm the presence and purity of the purified IspE homologs at each purification step (Fig. S2).

Gel filtration analysis: Purified samples of *EclspE*, *KplspE* and *AbispE* were prepared for the gel filtration analysis. To determine the molecular weights of our proteins, standard proteins of known molecular weights were used to calibrate the gel filtration column. Calibration standards included proteins ranging from 670 kDa to 1.35 kDa. A defined volume of the prepared *KplspE* sample was injected into the calibrated gel filtration column. The protein was detected using UV spectroscopy at wavelength of 280 nm. The elution profile of IspE was recorded based on the absorbance signals from the UV detector. The molecular weight of IspE was estimated by comparing its elution volume to the elution volumes of the calibration standards. A calibration curve was constructed using the known molecular weights of the standards (Fig. S2).

Enzymatic assay: To assess IspE activity and its kinetic parameters, we adapted a coupled spectrophotometric enzyme assay based on a previously described protocol (19). The analysis was carried out at room temperature (RT) over a 30-minute period. Enzymes and substrate concentrations were adapted for each homologue to ensure a suitable reaction velocity and linear kinetics. *EclspE* (200 nM), *KplspE* (150 nM) and *AbispE* (100 nM). For the kinetic assays, varying concentrations of the cofactor ATP and the substrate CDP-ME were employed. When one of these was held constant, ATP was maintained at 1 mM, and CDP-ME at 2 mM. For the determination of the IC₅₀ values of specific inhibitors, the percent of inhibition was calculated from a linear curve of initial velocities of the enzymes with different concentrations of the compounds. Kinetic parameters and IC₅₀ values of inhibitors were calculated using nonlinear regression analysis (OriginPro2023) (Fig. S3).

Crystallization: To determine the optimal crystallization condition of IspE homologues in the study, screening trials were carried out using commercially available crystallization kits from NeXtal. Plates were incubated at 18 °C and checked for crystals every three days. Repetitive optimizations were performed for the best identified crystal conditions. ***KplspE***: the structure was obtained after incubating 15 mg/mL *KplspE* with 5 mM of each of the AMP-PNP, CDP-ME and MgCl₂. Crystals were set in vapor diffusion chambers with a reservoir solution consisting of 1000 mM Sodium citrate, 100 mM HEPES pH 7.0. ***EclspE***: 7.5mg/mL protein was used to grow crystals in vapor diffusion chambers with a reservoir solution consisting of 25% PEG 4k; 200mM MgCl₂; 100mM MES pH 6.5. The crystals were then soaked for 2 days with 5–10 mM

compound. The best crystals were then mounted with 32% of glycerol (as a cryo-protectant) and data collection were carried out using beamtimes provided by the SLS and ESRF.

X-ray data collection and structures determination:

The structure of the new crystal form of *EclspE* and *KplspE* was determined through molecular replacement using Phaser.MR (38) within the Phenix software suite (39). The primary model used for the search was *EclspE* (PDB ID: 1OJ4) (14). Initially, data reduction and alignment were carried out using AIMLESS (40) within CCP4i (41). Following this, manual interpretation of the electron density maps was conducted using COOT (42). The 3D structures were refined iteratively using Phenix refine (43). Structures were validated in the PDB validation server before deposition. For visual representation, all figures were generated using PyMoL (43).

Comprehensive statistics regarding data collection and the refinement process can be located in Table S2.

Compounds synthesis

We sourced chemicals from commercial suppliers without further purification for our chemical synthesis. Chemical yield measurements are based on purified compounds, and were not optimized. Reaction progress was monitored with TLC silica gel sheets and UV visualization. Column chromatography used an automated CombiFlash® Rf system with RediSepRf silica columns. Preparative RP-HPLC was performed with a Thermo Fisher Scientific UltiMate 3000 system and a nucleodur® C18 Gravity column. ¹H and ¹³C NMR spectra were acquired using Bruker Avance Neo 500 MHz or Bruker Fourier 300 instruments. Chemical shifts were referenced against solvent peaks, and splitting patterns were noted. Low-resolution mass analysis and purity checks used an Ultimate 3000-MSQ LCMS system. Compounds for biological assays had ≥ 95% purity. High-resolution mass spectra were obtained using a ThermoFisher Scientific Q Exactive Focus system. Final products underwent thorough drying under high vacuum. For more detailed chemical and analytical methods, please refer to the supplementary materials.

Computational methods

Docking studies for IspE inhibitors to *EclspE* and *AblspE* were conducted using MOE v2020.09. The ligand preparation involved creating 2D structures of compounds and subjecting them to energy minimization. Protein structures were prepared for *EclspE* and *AblspE* through force-field application, hydrogen assignment, and energy minimization. Binding sites were selected based on proximity to ligands. Molecular docking was performed using MOE's Dock module, utilizing the co-crystallized ligand as a template and specific receptor structures. Scoring functions were applied to assess docking results.

Detailed computational methods can be found in the supplementary.

Acknowledgments

We acknowledge the European Synchrotron Radiation Facility (ESRF), Grenoble, France for provision of beamline ID23-1 and ID30B. We acknowledge the Paul Scherrer Institute, Villigen, Switzerland for the provision of synchrotron radiation beamtime at beamline X06DA-PXIII and X06SA-PXI of the SLS.

The authors would also like to thank Boris Illarionov for providing the substrate CDP-ME and Paula Kramer for helping with the computational work.

Author contributions A.K.H.H., R.H., conceived the study. R.H. wrote the main manuscript text and prepared all figures. D.W. and E.D. synthesized the compounds for bioassays. D.A. performed the bioassay, A.L. performed the docking study. M.M.H designed and supervised the medicinal chemistry work. All authors proofread the manuscript. A.K.H.H. supervised the project.

Funding The Helmholtz Association's Initiative and Networking Fund and the Schlumberger Foundation faculty for the future (FFTF) funded this work.

Competing interests The authors declare no competing interests.

Data availability The authors confirm that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work the author(s) used ChatGPT-3.5 in order to improve language and readability. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

5 References

1. World Health Organization (WHO) [online] <https://www.who.int> (Accessed August 8, 2023)
2. Pulingam, T., Parumasivam, T., Gazzali, A. M., Sulaiman, A. M., Chee, J. Y., Lakshmanan, M., Chin, C. F., and Sudesh, K. (2022) Antimicrobial resistance: Prevalence, economic burden, mechanisms of resistance and strategies to overcome. *Eur. J. Pharm. Sci.* **170**, 106103
3. Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993) Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295**, 517–524
4. Rohmer, M., Seemann, M., Horbach, S., Bringer-Meyer, S., and Sahm, H. (1996) Glyceraldehyde 3-Phosphate and Pyruvate as Precursors of Isoprenic Units in an Alternative Non-mevalonate Pathway for Terpenoid Biosynthesis. *J. Am. Chem. Soc.* **118**, 2564–2566
5. Illarionova, V., Kaiser, J., Ostrozhenkova, E., Bacher, A., Fischer, M., Eisenreich, W., and Rohdich, F. (2006) Nonmevalonate Terpene Biosynthesis Enzymes as Antiinfective Drug

- Targets: Substrate Synthesis and High-Throughput Screening Methods. *J. Org. Chem.* **71**, 8824–8834
6. Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000) Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13172–13177
 7. Gershenzon, J., and Dudareva, N. (2007) The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **3**, 408–414
 8. Heuston, S., Begley, M., Gahan, C. G. M., and Hill, C. (2012) Isoprenoid biosynthesis in bacterial pathogens. *Microbiol. Read. Engl.* **158**, 1389–1401
 9. Lell, B., Ruangweerayut, R., Wiesner, J., Missinou, M. A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H., and Kreamsner, P. G. (2003) Fosmidomycin, a Novel Chemotherapeutic Agent for Malaria. *Antimicrob. Agents Chemother.* **47**, 735–738
 10. Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H. K., Soldati, D., and Beck, E. (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science.* **285**, 1573–1576
 11. Lüttgen, H., Rohdich, F., Herz, S., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., Fellermeier, M., Sagner, S., Zenk, M. H., Bacher, A., and Eisenreich, W. (2000) Biosynthesis of terpenoids: YchB protein of *Escherichia coli* phosphorylates the 2-hydroxy group of 4-diphosphocytidyl-2C-methyl-d-erythritol. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1062–1067
 12. Kuzuyama, T., Takagi, M., Kaneda, K., Watanabe, H., Dairi, T., and Seto, H. (2000) Studies on the nonmevalonate pathway: conversion of 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol to its 2-phospho derivative by 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase. *Tetrahedron Lett.* **41**, 2925–2928
 13. Sgraja, T., Alphey, M. S., Ghilagaber, S., Marquez, R., Robertson, M. N., Hemmings, J. L., Lauw, S., Rohdich, F., Bacher, A., Eisenreich, W., Illarionova, V., and Hunter, W. N. (2008) Characterization of *Aquifex aeolicus* 4-diphosphocytidyl-2C-methyl-d-erythritol kinase – ligand recognition in a template for antimicrobial drug discovery. *FEBS J.* **275**, 2779–2794
 14. Miallau, L., Alphey, M. S., Kemp, L. E., Leonard, G. A., McSweeney, S. M., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F., and Hunter, W. N. (2003) Biosynthesis of isoprenoids: Crystal structure of 4-diphosphocytidyl-2C-methyl-d-erythritol kinase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9173–9178
 15. Wada, T., Kuzuyama, T., Satoh, S., Kuramitsu, S., Yokoyama, S., Unzai, S., Tame, J. R. H., and Park, S.-Y. (2003) Crystal Structure of 4-(Cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase, an Enzyme in the Non-mevalonate Pathway of Isoprenoid Synthesis*. *J. Biol. Chem.* **278**, 30022–30027
 16. Shan, S., Chen, X., Liu, T., Zhao, H., Rao, Z., and Lou, Z. (2011) Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*. *FASEB J.* **25**, 1577–1584
 17. Frank, A., and Groll, M. (2017) The Methylerythritol Phosphate Pathway to Isoprenoids. *Chem. Rev.* **117**, 5675–5703
 18. Kalinowska-Tłuścik, J., Miallau, L., Gabrielsen, M., Leonard, G. A., McSweeney, S. M., and Hunter, W. N. (2010) A triclinic crystal form of *Escherichia coli* 4-diphosphocytidyl-2 C -methyl- D -erythritol kinase and reassessment of the quaternary structure. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **66**, 237–241

19. Bernal, C., Mendez, E., Terencio, J., Boronat, A., and Imperial, S. (2005) A spectrophotometric assay for the determination of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase activity. *Anal. Biochem.* **340**, 245–251
20. Eoh, H., Narayanasamy, P., Brown, A. C., Parish, T., Brennan, P. J., and Crick, D. C. (2009) Expression and Characterization of Soluble 4-Diphosphocytidyl-2-C-methyl-D-erythritol Kinase from Bacterial Pathogens. *Chem. Biol.* **16**, 1230–1239
21. Shan, S., Chen, X., Liu, T., Zhao, H., Rao, Z., and Lou, Z. (2011) Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*. *FASEB J.* **25**, 1577–1584
22. H. Hirsch, A. K., S. Alphey, M., Lauw, S., Seet, M., Barandun, L., Eisenreich, W., Rohdich, F., N. Hunter, W., Bacher, A., and Diederich, F. (2008) Inhibitors of the kinase IspE: structure–activity relationships and co-crystal structure analysis. *Org. Biomol. Chem.* **6**, 2719–2730
23. Cheek, S., Zhang, H., and Grishin, N. V. (2002) Sequence and structure classification of kinases. *J. Mol. Biol.* **320**, 855–881
24. Krissinel, E., and Henrick, K. (2007) Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797
25. Patterson, H., Nibbs, R., McInnes, I., and Siebert, S. (2014) Protein kinase inhibitors in the treatment of inflammatory and autoimmune diseases. *Clin. Exp. Immunol.* **176**, 1–10
26. Haveliwala, D. D., Kamdar, N. R., Mistry, P. T., and Patel, S. K. (2014) Chromone-Fused Cytosine Analogues: Synthesis, Biological Activity, and Structure–Activity Relationship. *Nucleosides Nucleotides Nucleic Acids.* **33**, 80–91
27. Ropponen, H.-K., Diamanti, E., Johannsen, S., Illarionov, B., Hamid, R., Jaki, M., Sass, P., Fischer, M., Haupenthal, J., and Hirsch, A. K. H. Exploring the translational gap of a novel class of *Escherichia coli* IspE inhibitors. *ChemMedChem.* **n/a**, e202300346
28. Tang, M., Odejinmi, S., Allette, Y., Vankayalapati, H., and Lai, K. (2011) Identification of Novel Small Molecule Inhibitors of 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) kinase of Gram-negative bacteria. *Bioorg. Med. Chem.* **19**, 5886–5895
29. Crane, C. M., Hirsch, A. K. H., Alphey, M. S., Sgraja, T., Lauw, S., Illarionova, V., Rohdich, F., Eisenreich, W., Hunter, W. N., Bacher, A., and Diederich, F. (2008) Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis. *ChemMedChem.* **3**, 91–101
30. Mombelli, P., Le Chapelain, C., Munzinger, N., Joliat, E., Illarionov, B., Schweizer, W. B., Hirsch, A. K. H., Fischer, M., Bacher, A., and Diederich, F. (2013) Imidazole- and Benzimidazole-Based Inhibitors of the Kinase IspE: Targeting the Substrate-Binding Site and the Triphosphate-Binding Loop of the ATP Site: Imidazole- and Benzimidazole-Based Inhibitors of the Kinase IspE. *Eur. J. Org. Chem.* **2013**, 1068–1079
31. Hirsch, A. K. H., Lauw, S., Gersbach, P., Schweizer, W. B., Rohdich, F., Eisenreich, W., Bacher, A., and Diederich, F. (2007) Nonphosphate Inhibitors of IspE Protein, a Kinase in the Non-Mevalonate Pathway for Isoprenoid Biosynthesis and a Potential Target for Antimalarial Therapy. *ChemMedChem.* **2**, 806–810
32. Tidten-Luksch, N., Grimaldi, R., Torrie, L. S., Frearson, J. A., Hunter, W. N., and Brenk, R. (2012) IspE Inhibitors Identified by a Combination of In Silico and In Vitro High-Throughput Screening. *PLoS ONE.* **7**, e35792
33. Persch, E., Dumele, O., and Diederich, F. (2015) Molecular Recognition in Chemical and Biological Systems. *Angew. Chem. Int. Ed.* **54**, 3290–3327
34. Hendsch, Z. S., and Tidor, B. (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci. Publ. Protein Soc.* **3**, 211–226

35. Mecozzi, S., and Rebek, J., Jr. (1998) The 55 % Solution: A Formula for Molecular Recognition in the Liquid State. *Chem. – Eur. J.* **4**, 1016–1022
36. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303
37. Molecular Operating Environment (MOE), 2022.02 Chemical Computing Group ULC, 910-1010 Sherbrooke St. W., Montreal, QC H3A 2R7, Canada, 2023.
38. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
39. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
40. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214
41. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
42. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
43. Yuan, S., Chan, H. C. S., Filipek, S., and Vogel, H. (2016) PyMOL and Inkscape Bridge the Data and the Data Visualization. *Struct. Lond. Engl. 1993.* **24**, 2041–2042

Supplementary:**IspE Kinase Homologues: Characterization and crystal Structures, enable structure–based design of novel anti-infectives**

Rawia Hamid^{1, 2}, Danica J. Walsh¹, Eleonora Diamanti¹, Diana Aguilar¹, Antoine Lacour^{1, 2}, Mostafa M. Hamed¹, Anna K. H. Hirsch^{1, 2*}.

1. Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) – Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123 Saarbrücken, Germany.

2. Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany.

*Correspondence e-mail: anna.hirsch@helmholtz-hips.de

Submitted manuscript

Figure S1: Sequence alignment of the IspE from <i>K. pneumoniae</i> , <i>E. coli</i> , <i>A. baumannii</i>	177
Figure S2: Gel filtration calibration curve used to estimate molecular weights	178
Figure S3: Michaelis-Menten kinetics analysis, K_m values of <i>EclspE</i> , <i>KplspE</i> and <i>AbIspE</i>	179
Figure S3: Inhibitory dose-response curves	180
Table S1: Data collection and refinement statistics	181
Scheme S1: Compounds synthesis	182
Table S2: Molecular docking of compounds to <i>AbIspE</i> structure	197

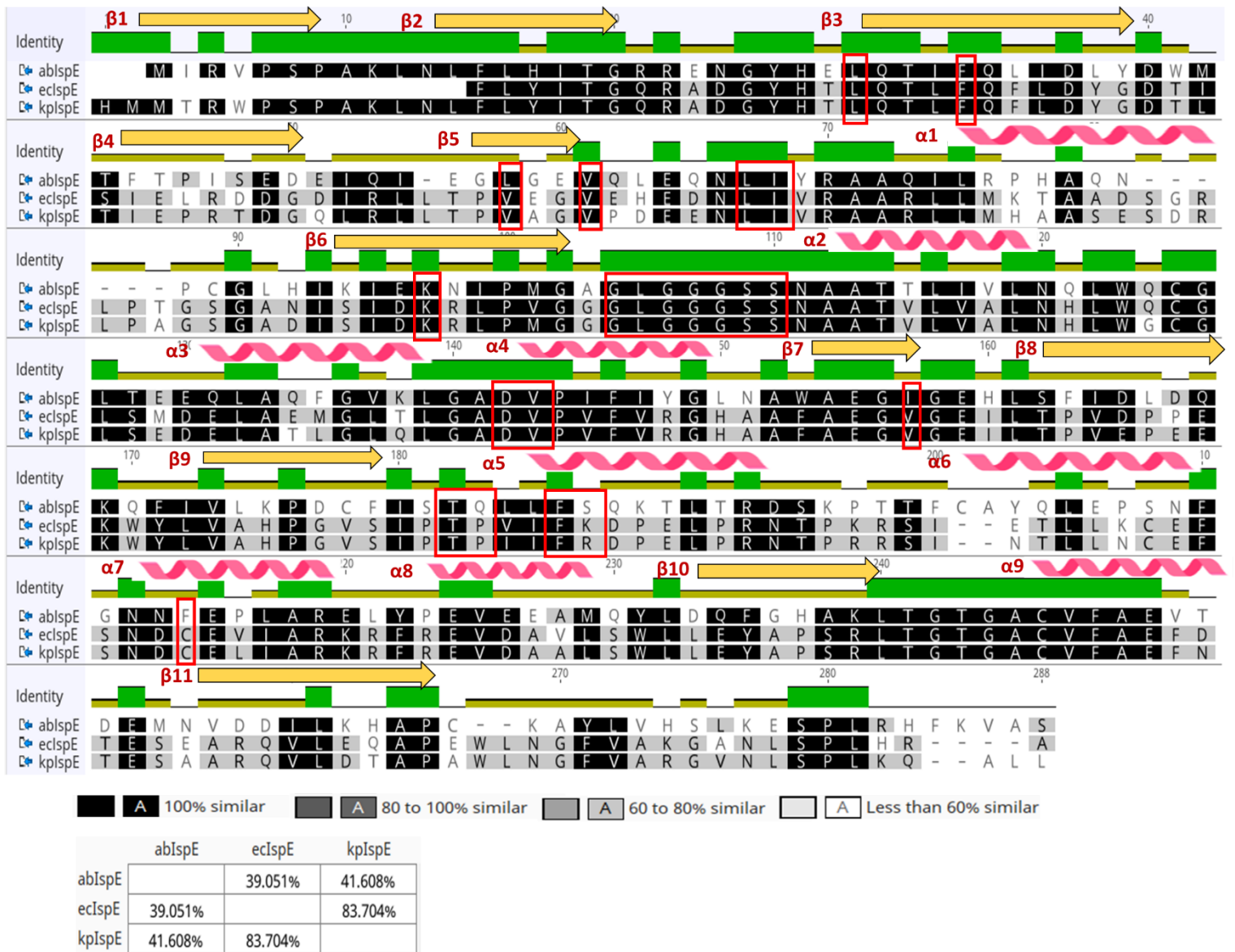


Figure S1: Sequence alignment of the IspE from *K. pneumoniae*, *E. coli*, *A. baumannii*. Secondary structure elements of IspE are indicated at top of the alignment; α helices and β strands are presented as curves and arrows. The identity, shown as a bar graph above the sequences, was calculated using the three aligned sequences. Amino acids in the active site are indicated by red boxes.

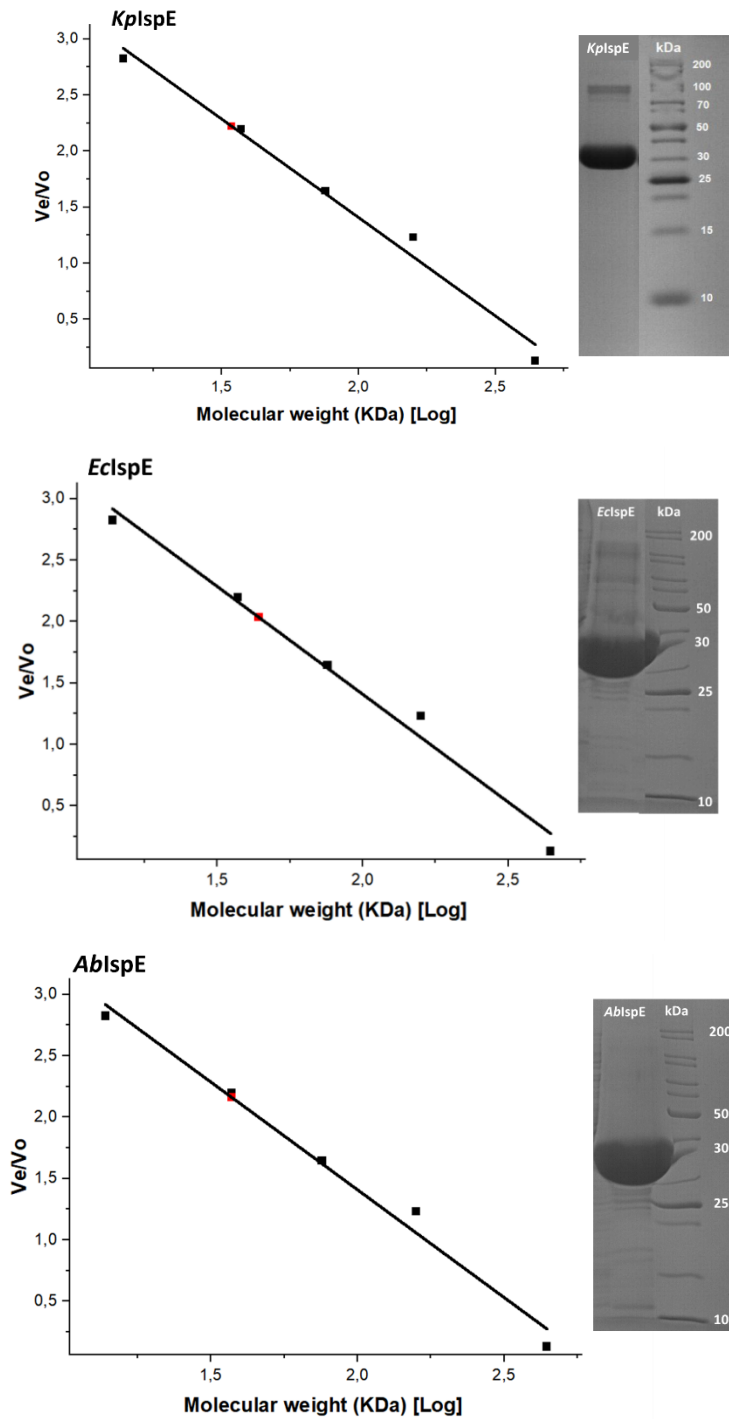


Figure S2: Gel filtration calibration curve used to estimate molecular weights for *KplspE*, *EclspE* and *AblspE*. Size exclusion chromatography column S200 16/600 was calibrated using Thyroglobulin (670 KDa), γ -Globulin (158 KDa), ovalbumin (44 kDa), Myoglobin (17 KDa), and Vitamin B12 (1,350 KDa). The calibration curve was plotted using the elution volume/ column void volume (V_e/V_0) versus logarithm of the molecular weight. Straight line is the calibration curve calculated from the data for molecular weight standards. Red dots correspond to the positions of V_e/V_0 values for *KplspE*, *EclspE* and *AblspE* respectively. Linear equation, from the calibration curve was used to calculate the experimental molecular weights of *KplspE* (31 kDa), *EclspE* (34 kDa) and *AblspE* (30 kDa). SDS-PAGE 12% electrophoresis of IspEs showing the expected band size obtained purification.

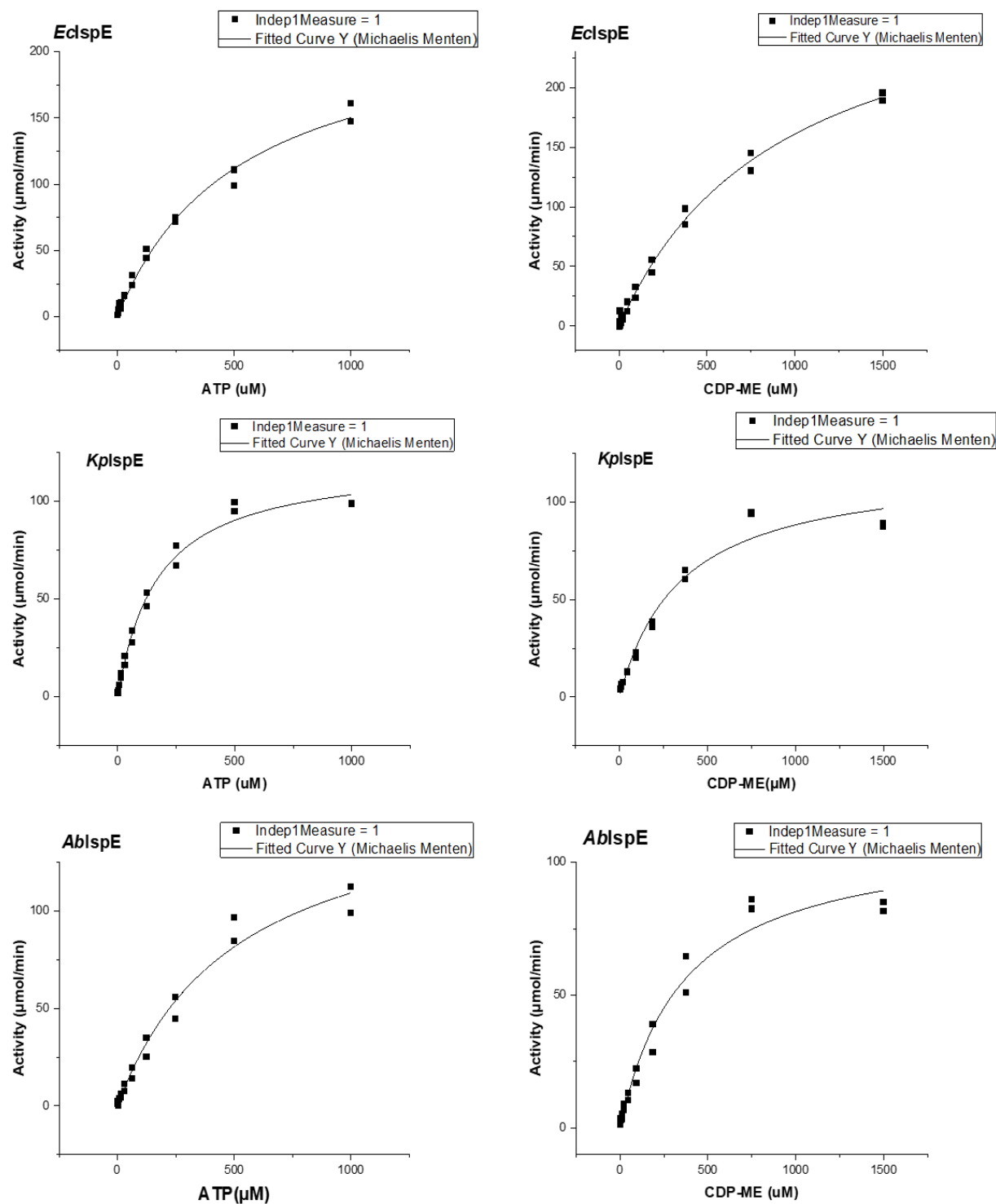


Figure S3: Michaelis-Menten kinetics analysis, K_m values of *EclspE*, *KplspE* and *AblspE* enzymes, analysis was conducted by varying concentrations of ATP and CDP-ME, enzyme concentration used: *EclspE* (200nM), *KplspE* 150nM) and *AblspE* (100nM). When a substrate or cofactor was kept constant, a concentration of 1mM was used for ATP and 1.5 mM for CDP-ME.

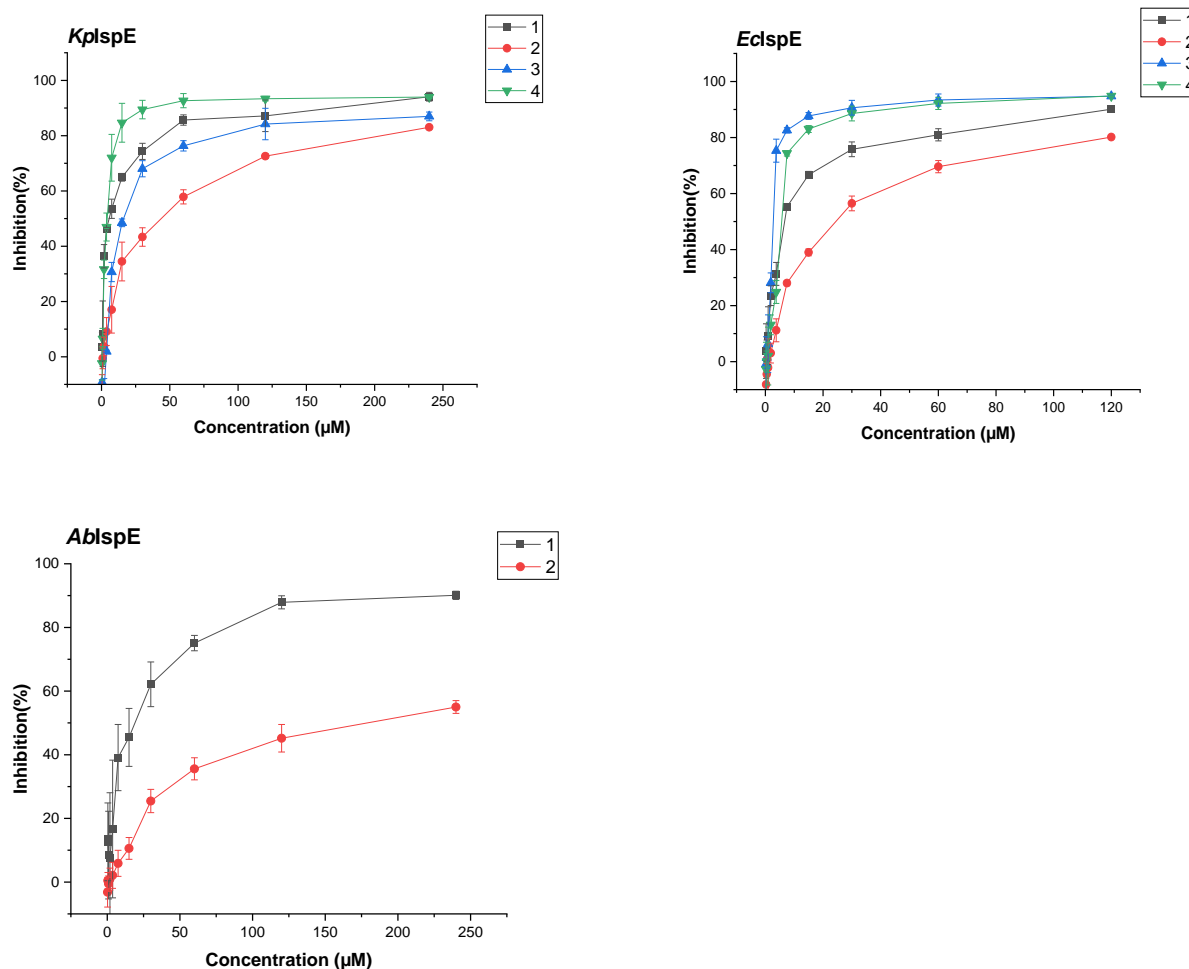


Figure S4: Inhibitory dose-response curves to determine the IC_{50} for each inhibitor on *EclspE*, *KplspE* and *AblspE*. The curves were plotted using % of inhibition values of the enzymes after incubation with varied concentration of compounds. *EclspE* concentration (200nM), *KplspE* (150nM) and *AblspE* (100nM). CDP-ME was used at ($\sim 1X K_m$); 300 μM for *EclspE* and *KplspE* assay and 400 μM for *AblspE*. ATP was kept at 1mM ($\sim 4X K_m$) in all instances. The curves were fit using non-linear regression analysis program in OriginPro.

Table S1: Data collection and refinement statistics of all X-ray structures shown in the manuscript. Statistics for the highest-resolution shell are shown in parentheses

	<i>KplspE</i>	<i>EclspE</i>	<i>EclspE</i> (1)	<i>EclspE</i> (2)	<i>EclspE</i> (3)
PDB code	8CKH	8QC7	8QCC	8QCN	8QCO
Resolution range	41.46 - 1.8	32.76 - 2.03	47.93 - 1.904	36.43 - 1.86	48.28 - 1.55
Space group	P21 21 21	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell a b c [Å] α β γ [°]	65.87 76.781 98.517 90 90 90	143.196 53.046 91.07 90 126.933 90	142.57 52.793 90.253 90 126.691 90	143.928 53.1129 91.7173 90 127.395 90	145.674 53.162 91.968 90 127.677 90
Total reflections		35245	41772	39594	79611
Unique reflections	46367 (4526)	35241 (3472)	41807 (3718)	45832 (4238)	79873 (7952)
Multiplicity	6.3	3.9	10.2	2.4	2.8
Wavelength	0.97	0.93	0.93	0.93	0.62
Completeness (%)	98.9 (98.1)	98.97 (98.52)	98.41 (90.28)	95.7	98.23 (98.82)
Mean I/ sigma (I)	2.09 (at 1.79Å)	2.4 (at 2.0Å)	2.0 (at 1.9 Å)	7.6 (at 1.86 Å)	10.4 (1.55 Å)
Wilson B-factor	26.9	26.81	37.59	39.28	26.21
R-merge	0.079	0.12	0.078	0.058	0.046
R-work	0.1885 (0.2348)	0.2165 (0.3117)	0.1935 (0.3676)	0.2271 (0.5133)	0.2104 (0.3619)
R-free	0.2267 (0.2870)	0.2404 (0.3794)	0.2353 (0.3785)	0.2660 (0.5499)	0.2411 (0.4087)
Number of non-hydrogen atoms	4633	4627	4605	4523	4707
Macromolecules	4275	4334	4314	4330	4332
Ligands	88	192	165	138	150
Solvent	296	171	186	103	281
Protein residues	559	562	561	563	563
RMS (bonds)	0.005	0.020	0.041	0.024	0.031
RMS (angles)	0.80	0.75	1.27	0.66	1.17
Average B-factor	33.65	33.24	41.86	48.57	33.14
Ramachandran plot analysis (%)					
Favored	98.38	99.1	98.56	97.85	99.11
Allowed	1.62	0.90	1.44	2.15	0.89
Outliers	0.00	0.00	0.00	0.00	0.00

Compounds synthesis:

All reagents used for chemical synthesis were purchased from commercial suppliers, and used without further purification. All chemical yields refer to purified compounds and were not optimized. Reaction progress was monitored using TLC silica gel 60 F₂₅₄ aluminum sheets, and visualization was accomplished by UV at 254 nm. Column chromatography was performed using the automated flash chromatography system CombiFlash® Rf (Teledyne Isco) equipped with RediSepRf silica columns. Preparative RP-HPLC was performed using an UltiMate 3000 Semi-Preparative System (Thermo Fisher Scientific) with nucleodur® C18 Gravity (250 mm × 10 mm, 5 μm) column. Purifications *via* preparative RP-HPLC were carried out using two possible conditions: *Condition A*: gradient 5–100% CH₃CN + 0.05% HCOOH in water + 0.05% HCOOH in 53 min at a flow rate of 5 mL/min. The sample was dissolved in DMSO and manually injected to the HPLC system. *Condition B*: gradient 20–100% CH₃CN + 0.05% HCOOH in water + 0.05% HCOOH in 35 min at a flow rate of 5 mL/min. Samples were dissolved in DMSO and manually injected to the HPLC system. ¹H and ¹³C NMR spectra were recorded as indicated on a Bruker Avance Neo 500 MHz (¹H, 500 MHz; ¹³C, 126 MHz) with prodigy cryoprobe system or a Bruker Fourier 300 (¹H, 300 MHz; ¹³C, 75 MHz) instrument. Chemical shifts were recorded as δ values in ppm units and referenced against the residual solvent peak (CDCl₃, δ = 7.26, 77.16; DMSO-d₆, δ = 2.50, 39.52, MeOH-d₄: δ = 3.35, 4.78, 49.3). Splitting patterns describe apparent multiplicities and are designated as s (singlet), br s (broad singlet), d (doublet), dd (doublet of doublet), t (triplet), q (quartet), m (multiplet). Coupling constants (*J*) are given in hertz (Hz). Low resolution mass analytics and purity control of final compounds was carried out either using an Ultimate 3000-MSQ LCMS system (Thermo Fisher Scientific) consisting of a pump, an autosampler, MWD detector and a ESI quadrupole mass spectrometer. Purity of all compounds used in biological assays was ≥ 95%. High resolution mass spectra were recorded on a ThermoFisher Scientific (TF, Dreieich, Germany) Q Exactive Focus system equipped with heated electrospray ionization (HESI)-II source. Final products were dried at high vacuum.

General Procedure A for the formation of a sulfonamide:

The sulfonyl chloride (1 eq.) was added dropwise to a solution of propargyl amine (1 eq.) and Et₃N (1.1 eq.) in dry CH₂Cl₂, at 0 °C. The mixture was left to stir at 25 °C for 30 min and concentrated *in vacuo*. Products were purified via column chromatography.

General procedure B for the Sonogashira cross coupling:

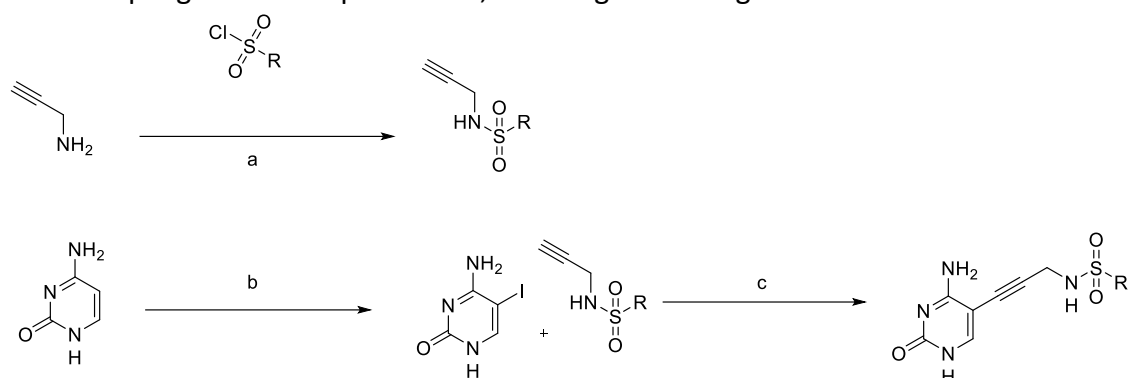
To an Ar-charged flask charged with iodocytosine (1.0 eq.), the previously synthesized alkyne (2 eq.), and Et₃N (3.0 eq.) in anhydrous DMF, [PdCl₂(PPh₃)₂] (0.1 eq.) and CuI (0.2 eq.) were added at 25 °C. The mixture was left to stir at 25 °C for 12 h. Water was added to the flask and the reaction mixture filtered, dissolved in DMSO and filtered again through celite. The reaction mixture was purified via HPLC.

Chemical and Analytical Methods.

¹H- and ¹³C-NMR spectra were recorded as indicated on a Bruker Avance Neo 500 MHz with prodigy cryoprobe system or a Bruker Fourier 300 instrument. Chemical shifts are given in parts per million (ppm), and referenced against the residual solvent peak. Coupling constants (*J*) are given in Hertz. Low-resolution mass analytics and purity control of final compounds were carried out either using an Ultimate 3000-MSQ LCMS system (Thermo Fisher Scientific), consisting of a pump, an autosampler, MWD detector and an ESI quadrupole mass spectrometer. High resolution mass spectra were recorded on a ThermoFisher Scientific (TF, Dreieich, Germany) Q Exactive Focus system equipped with heated electrospray ionization (HESI)-II source. Reagents were used as obtained from commercial suppliers without further

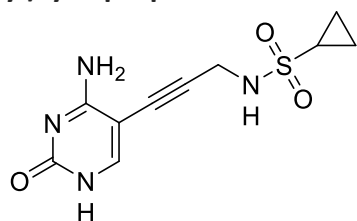
purification. Procedures were not optimized regarding yield. Column chromatography was performed using the automated flash chromatography system CombiFlash® Rf (Teledyne Iso) equipped with RediSepRf silica columns. Preparative HPLC was performed using an UltiMate 3000 Semi-Preparative System (Thermo Fisher Scientific) with nucleodur® C18 Gravity (250 mm × 10 mm, 5 µm) column. Final products were dried at high vacuum.

Derivatives of **1** were synthesized in three steps (Scheme S 1). First, sulfonamides were synthesized from the corresponding sulfonyl chloride and propargylamine. Iodination of cytosine was performed using I₂ and iodic acid to yield 5-iodocytosine. Finally, a sonogashira cross coupling reaction is performed, resulting in the target derivative.



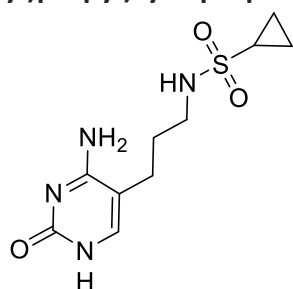
Scheme S 1: Reagents and condition: a) TEA, DCM, 0°C-r.t, 30min; b) I₂, HIO₃, AcOH, 40°C, 12h; c) [PdCl₂(PPh₃)₂], CuI, TEA, DMF, r.t, 4-12 h;

Synthesis of N-(3-(4-amino-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)cyclopropanesulfonamide (**2**)



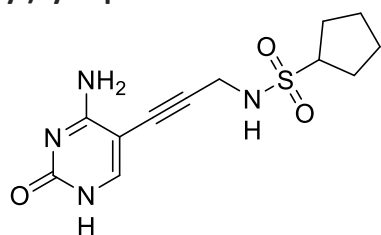
The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition B). The title product was obtained as a white solid, (70 mg, 0.26 mmol, 38%). ¹H NMR (500 MHz, (DMSO-d₆) δ 8.52 (s, 1H), 4.16 (s, 2H), 1.56 (m, 1H), 0.69 (m, 4H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 165.87, 155.54, 146.56, 142.20, 91.79, 76.12, 33.45, 30.07, 5.49. Purity: 100%; MS (ESI+) *m/z* 287 [M+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₀H₁₂N₄O₃S [M+H]⁺: 287.0630 found 287.0424.

Synthesis of N-(3-(4-amino-2-oxo-1,2-dihydropyrimidin-5-yl)propyl)cyclopropanesulfonamide



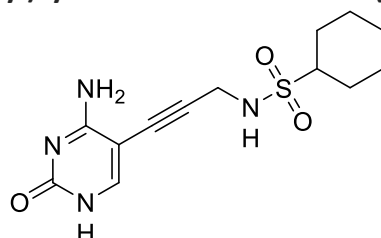
To a stirring solution of 1 (100 mg, 0.38 mmol, 1 eq.) in EtOH, Pd/C (10 mg, 0.1 mmol, 0.25 eq.) was added. H₂ gas was bubbled in to the stirring solution and the reaction mixture was allowed to stir for 3 h. The mixture was filtered over celite and washed with EtOH and concentrated under reduced pressure. The title product was obtained as a white solid, (92 mg, 0.34 mmol, 89%). ¹H NMR (500 MHz, (DMSO-d₆) δ 7.24 (s, 1H), 3.05 (t, *J*=6.1 Hz, 2H), 3.05 (t, *J*=6.1 Hz, 2H), 2.59-2.57 (m, 1H), 2.33 (t, *J*=7.2 Hz, 2H), 1.66 (tt, *J*=7.4 Hz, 2H), 1.00-0.94 (m, 4H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 157.39, 146.65, 142.28, 141.63, 44.47, 33.39, 27.99, 26.25, 13.71. Purity: 95.51%; MS (ESI+) *m/z* 273 [*M*+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₀H₁₆N₄O₃S [*M*+H]⁺: 273.0943 found 273.1013.

Synthesis of N-(3-(4-amino-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)cyclopentanesulfonamide (3)



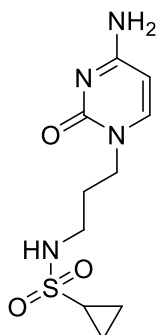
The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition B). The title product was obtained as a white solid, (74 mg, 0.25 mmol, 36%). ¹H NMR (500 MHz, (DMSO-d₆) δ 7.67 (s, 1H), 4.03 (d, *J*= 5.6 HZ, 2H), 3.68-3.64 (m, 1H), 1.93-1.84 (m, 2H), 1.68-1.53 (m, 2H), 1.28-1.18 (m, 4H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 167.59, 160.12, 140.16, 134.79, 87.08, 80.27, 54.41, 53.71, 33.34, 23.00. Purity: 100%; MS (ESI+) *m/z* 297 [*M*+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₂H₁₆N₄O₃S [*M*+H]⁺: 297.0943 found 297.1021.

Synthesis of N-(3-(4-amino-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)cyclohexanesulfonamide (4)



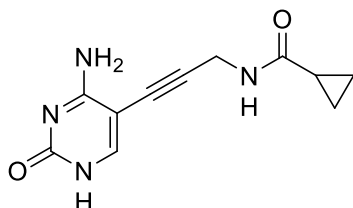
The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition B). The title product was obtained as a white solid, (8 mg, 0.026 mmol, 28%). ¹H NMR (500 MHz, (DMSO-d₆) δ 7.67 (s, 1H), 4.00 (s, 2H), 3.07 (t, *J*= 8.7 Hz, 1H), 2.07-2.06 (m, 1H), 1.78-1.75 (m, 2H), 1.62-1.59 (m, 1H), 1.37-1.09 (m, 6H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 159.03, 153.85, 141.32, 128.22, 48.49, 47.03, 31.41, 26.80, 23.90. Purity: 99.8%; MS (ESI+) *m/z* 311 [*M*+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₃H₁₈N₄O₃S [*M*+H]⁺: 311.1100 found 311.1770.

Synthesis of N-(3-(4-amino-2-oxopyrimidin-1(2H)-yl)propyl)cyclopropanesulfonamide (HIPS7996)



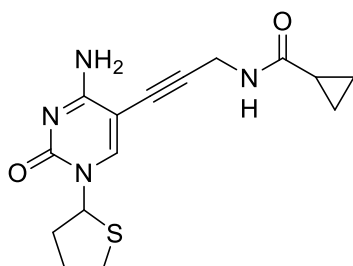
A solution of TEA (0.1 mL, 0.81 mmol, 1.2 eq) and of 4-amino-1-(3-aminopropyl) pyrimidin-2(1H)-one (115 mg, 0.68 mmol, 1 eq.) in DCM was cooled to 0 °C. Cyclopropanesulfonyl chloride (0.54 mL, 0.54 mmol, 0.8 eq.) was added to the solution with stirring. The reaction was allowed to warm to room temperature and stir for an addition 2 hours. The reaction mixture was concentrated and purified via HPLC (condition A). The title product was obtained as a white solid (14 mg, 0.05 mmol, 8%). ¹H NMR (500 MHz, (DMSO-d₆) δ 7.98 (d, *J*= 7.6 Hz, 1H), 6.03 (d, *J*= 7.4 Hz, 1H), 3.82 (t, *J*= 6.8 Hz, 2H), 3.03 (q, *J*= 6.2, 6.5 Hz, 2H), 1.85-1.82 (m, 1H), 1.19 (t, *J*= 7.3 Hz, 2H), 0.95-0.89 (m, 2H), 0.57-0.50 (m, 2H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 173.27, 162.36, 161.59, 125.01, 98.26, 96.90, 40.92, 29.53, 28.75, 10.20. Purity: 100%; MS (ESI+) *m/z* 272, [M+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₀H₁₆N₄O₃S [M+H]⁺: 273.3230 found 273.2943.

Synthesis of N-(3-(4-amino-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)cyclopropanecarboxamide (HIPS7876)



The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition A). The title product was obtained as a white solid, (73 mg, 0.3 mmol, 23%). ¹H NMR (500 MHz, (DMSO-d₆) δ 7.67 (s, 1H), 4.04 (s, 2H), 2.66-2.61 (m, 1H), 0.97-0.95 (m, 4H). ¹³C NMR (126 MHz, (DMSO-d₆) δ 177.14, 155.40, 150.95, 136.22, 129.64, 85.37, 82.71, 32.81, 15.69, 9.37. Purity via HPLC: 100%; MS (ESI+) *m/z* 233 [M+H]⁺; HRMS (ESI+) *m/z* calcd for C₁₁H₁₂N₄O₂ [M+H]⁺: 233.0960 found 233.1031.

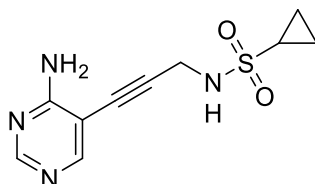
N-(3-(4-amino-2-oxo-1-(tetrahydrothiophen-2-yl)-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)cyclopropanecarboxamide (HIPS8061)



The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition A). The title product was obtained as a white solid. (20 mg, 0.06 mmol, 10%). ¹H NMR (500 MHz, (DMSO-d₆) δ 8.18 (s, 1H), 6.22 (dd, *J*= 5.6, 5.8 Hz, 1H), 4.18 (d, *J*= 5.0 Hz, 2H), 3.32 (ddd, *J*= 6.4, 6.0, 4.1 Hz, 1H), 2.91 (ddd, *J*= 6.4, 6.4, 3.7 Hz,

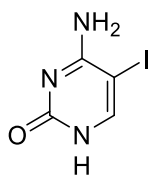
1H), 2.28-2.23 (m, 1H), 2.13-2.01 (m, 3H), 1.63 (tt, $J = 5.3, 5.4$ Hz, 1H), 0.75-0.73 (m, 4 H). ^{13}C NMR (126 MHz, (DMSO- d_6) δ 172.47, 163.52, 161.33, 130.52, 126.88, 88.54, 83.51, 64.23, 44.07, 32.57, 28.65, 23.78, 15.84, 6.72. Purity via HPLC: 100%; MS (ESI+) m/z 319 [$M+H$] $^+$; HRMS (ESI+) m/z calcd for $\text{C}_{15}\text{H}_{18}\text{N}_4\text{O}_2\text{S}$ [$M+H$] $^+$: 319.1150 found 319.1218.

Synthesis of N-(3-(4-aminopyrimidin-5-yl)prop-2-yn-1-yl)cyclopropanesulfonamide (HIPS8013)



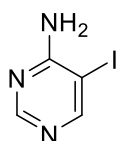
The title compound was synthesized according to the general method B. The crude product was purified by preparative HPLC (Condition B). The title product was obtained as a white solid, (5 mg, 0.02 mmol, 21%). ^1H NMR (500 MHz, (DMSO- d_6) δ 8.08 (d, $J = 7.2$ Hz, 1H), 7.14 (d, $J = 7.3$ Hz, 1H), 3.80 (t, $J = 7.1$ Hz, 2H), 3.09 (q, $J = 6.8, 6.1$ Hz, 2H), 1.78 (t, $J = 6.9$ Hz, 2H), 1.51-1.49 (m, 1H), 0.87-0.84 (m, 2H), 0.67-0.63 (m, 2H). ^{13}C NMR (126 MHz, (DMSO- d_6) δ 154.91, 150.40, 129.30, 93.72, 47.34, 46.13, 29.50, 28.93, 5.08, 4.72. Purity: 100%; MS (ESI+) m/z 252, [$M+H$] $^+$; HRMS (ESI+) m/z calcd for $\text{C}_{10}\text{H}_{12}\text{N}_4\text{O}_2\text{S}$ [$M+H$] $^+$: 253.2920 found 253.0736.

Synthesis of 6-Amino-5-iodo-2(1H)-pyrimidinone (5-iodocytosine)



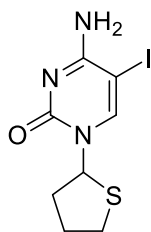
Cytosine (1 g, 9 mmol, 1 eq.), iodine (3.2 g, 13.5 mmol, 1.5 eq.) and iodic acid (2.2 g, 12.6 mmol, 1.4 eq.) were stirred in acetic acid at 40 °C overnight. The reaction mixture was cooled and treated with sat. aq. $\text{Na}_2\text{S}_2\text{O}_3$ until a white suspension was obtained. The mixture was neutralized with NaOH (10%). The product was collected via filtration, washed with cold water (5x) and concentrated under reduced pressure. $_$ was obtained as a white solid. Yield 97%. ^1H NMR (500 MHz, (CDCl $_3$) δ 7.98 (s, 1H). MS (ESI+) m/z 237.

Synthesis of 5-iodopyrimidin-4-amine



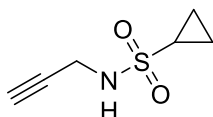
4-Aminopyrimidine (0.5 g, 5.25 mmol, 1 eq.), N-iodosuccinimide (1.2 mL, 12.6 mmol, 2.4 eq.) were stirred in acetic acid at 80 °C with stirring for 2.5 hours. The reaction mixture was cooled and extracted with DCM (3x). The organic layer was washed with sat. aq. $\text{Na}_2\text{S}_2\text{O}_3$ (3x) and brine (3x) and then concentrated under reduced pressure. The title product was obtained as a yellow solid. Yield 91%. ^1H NMR (500 MHz, (CDCl $_3$) δ 8.44 (s, 1H), 8.30 (s, 1H).

4-amino-5-iodo-1-(tetrahydrothiophen-2-yl)pyrimidin-2(1H)-one



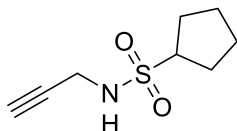
To a suspension of Synthesis of 5-iodopyrimidin-4-amine (0.30 g, 1.26 mmol, 1.0 eq) in a mixture toluene/DMF (2:1, 12 mL), triethylamine (0.70 mL, 5.06 mmol, 4 eq) and TMSOTf (1.8 mL, 10.08 mmol, 8eq) were added. The mixture was stirred at rt for 15 min. Then, tetrahydrothiophene 1-oxide (0.11 mL, 1.26 mmol, 1 eq) was added dropwise followed by additional triethylamine (1.05 mL, 7.56 mmol, 6eq). The mixture was stirred at rt for 30 min. The reaction was quenched by the addition of ice and extracted with EtOAc, followed by being washed with sat. aq. NaHCO₃ (3 x 10 mL). The crude product was purified by preparative HPLC (Condition B). The title compound was obtained as a white solid. Yield: 12 %. ¹H NMR (500 MHz, (DMSO-d₆) δ 8.15 (s, 1H), 6.12 (dd, *J*= 6.4 Hz, 1H), 3.29-3.22 (m, 1H), 2.87-2.83 (m, 1H), 2.20-2.02 (m, 1H), 2.06-1.96 (m, 3H).

Synthesis of N-(prop-2-yn-1-yl)cyclopropanesulfonamide



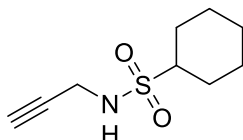
The title compound was synthesized according to the general method A. The crude product was purified by column chromatography using a gradient of 0-100% EtOAc in Hexanes to yield compounds. Yield 69%. ¹H NMR (500 MHz, (CDCl₃) δ 3.91 (dd, *J*= 2.5, 3.7 Hz, 2H), 2.51-2.46 (m, 1H), 2.28 (t, *J*= 2.4 Hz, 1H), 1.18-1.15 (m, 2H), 1.00-0.96 (m, 2H). MS (ESI+) *m/z* 160.

Synthesis of N-(prop-2-yn-1-yl)cyclobutanesulfonamide



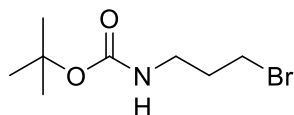
The title compound was synthesized according to the general method A. The crude product was used without any further purification.

Synthesis of N-(prop-2-yn-1-yl)cyclohexanesulfonamide



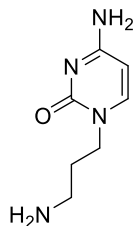
The title compound was synthesized according to the general method A. The crude product was purified by column chromatography using a gradient of 0-100% EtOAc in Hexanes to yield compounds. Yield 35%; ¹H NMR (500 MHz, (CDCl₃) δ 3.88 (dd, *J*= 2.5, 3.7 Hz, 2H), 2.97-2.92 (m, 1H), 2.27 (s, 1H), 2.17-2.14 (m, 2H), 1.86-1.83 (m, 2H), 1.66-1.63 (m, 1H), 1.49-1.44 (m, 5H), 1.26-1.12 (m, 2H); MS (ESI+) *m/z* 202.

Synthesis of tert-butyl (3-bromopropyl)carbamate



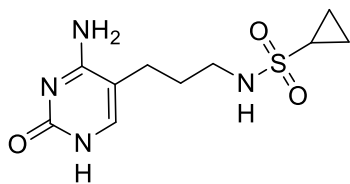
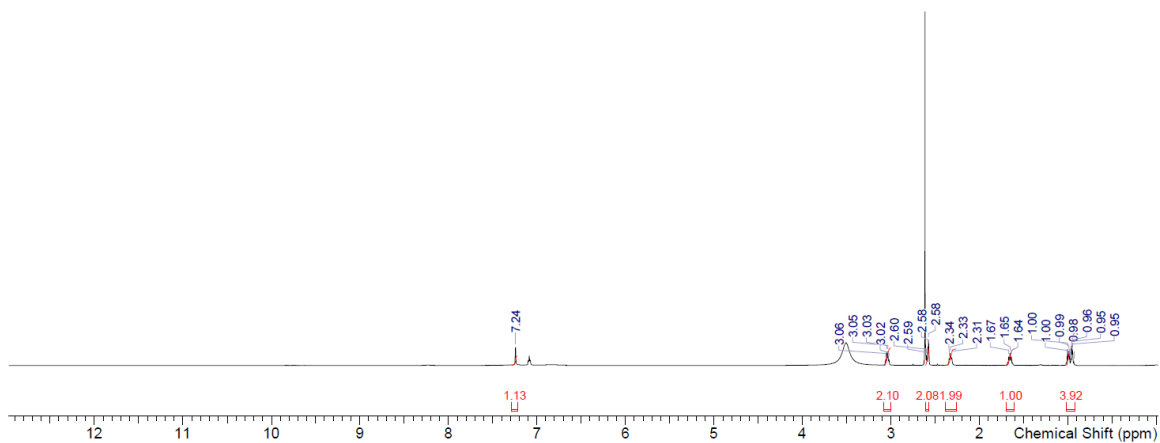
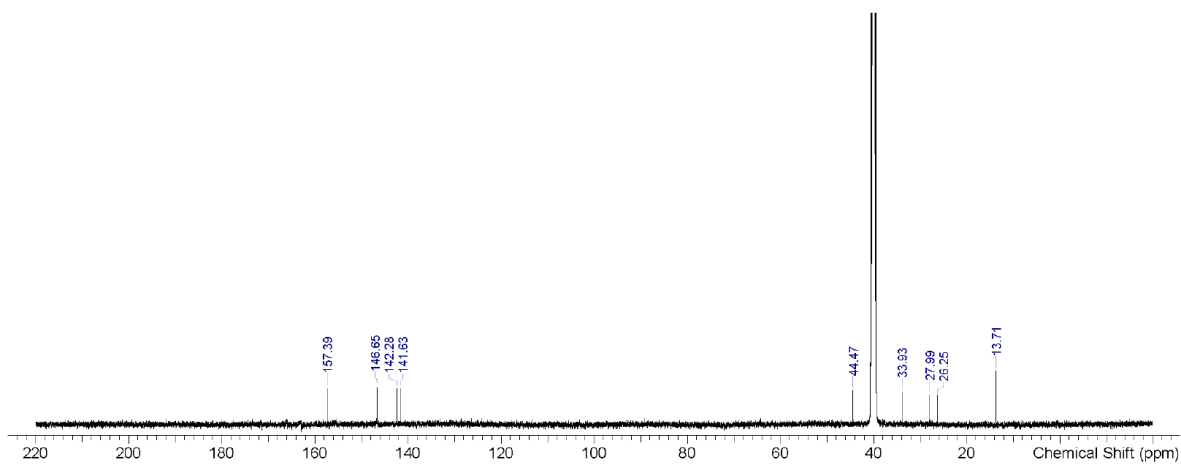
A mixture of 3-bromopropylamin hydrobromide (2 g, 9.14 mmol, 1 eq.), (BOC)₂O (2.1 g, 9.59 mmol, 1.05 eq.) and triethyl amine (6.5 ml, 29.3 mmol, 3.2 eq.) was stirred in dry DCM at room temperature for 24 h. The crude mixture was filtered and washed with diethyl ether to yield the product as a white solid. Yield 89%. ¹H NMR (500 MHz, (CDCl₃) δ 3.56 (t, *J*= 6.7 Hz, 2H), 3.09 (t, *J*= 6.4 Hz, 2H), 1.97 (q, *J*= 6.5, 6.7 Hz, 2H), 1.44 (s, 9H).

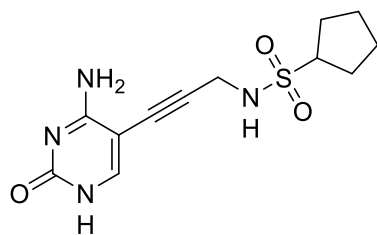
Synthesis of 4-amino-1-(3-aminopropyl)pyrimidin-2(1H)-one



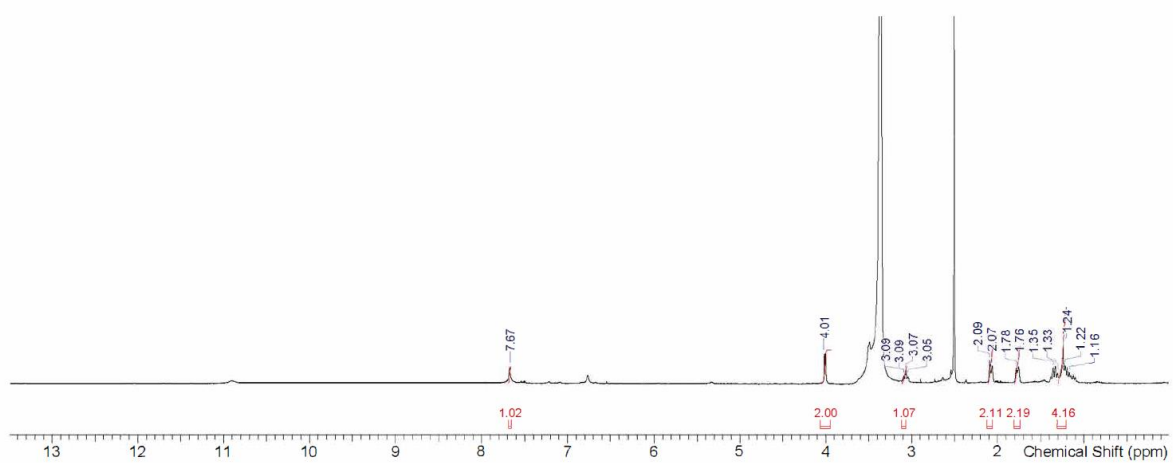
Cytosine (388 mg, 3.5 mmol, 1.2 eq.) and 60% NaH (147mg, 3.7 mmol, 1.05 eq.) were dissolved in dry DMF and stirred for 10 minutes. BOC-3-bromopropylamine (1 g, 4.2 mmol, 1 eq.) was added slowly and the reaction mixture was stirred at room temperature for 4 h. Water was added and the mixture was extracted with DCM (3x). The organic layer was concentrated in vacuum to obtain a yellow oil. The crude product was then dissolved in a methanol and acidified with HCl. The reaction mixture was stirred at room temperature for 1.5 hours. The reaction was concentrated under vacuum and dissolved in DCM. The crude mixture was washed with water (3x), brine (3x) and then purified via column chromatography. Yield 46%. ¹H NMR (500 MHz, (DMSO-d₆) δ 7.50 (d, *J*= 7.0, 1H), 5.56 (d, *J*= 7.1, 1H), 3.53 (t, *J*= 6.9 Hz, 2H), 3.03 (q, *J*= 6.6, 6.6.1 Hz, 2H), 1.58 (t, *J*= 6.7 Hz, 2H).

NMR Spectra:

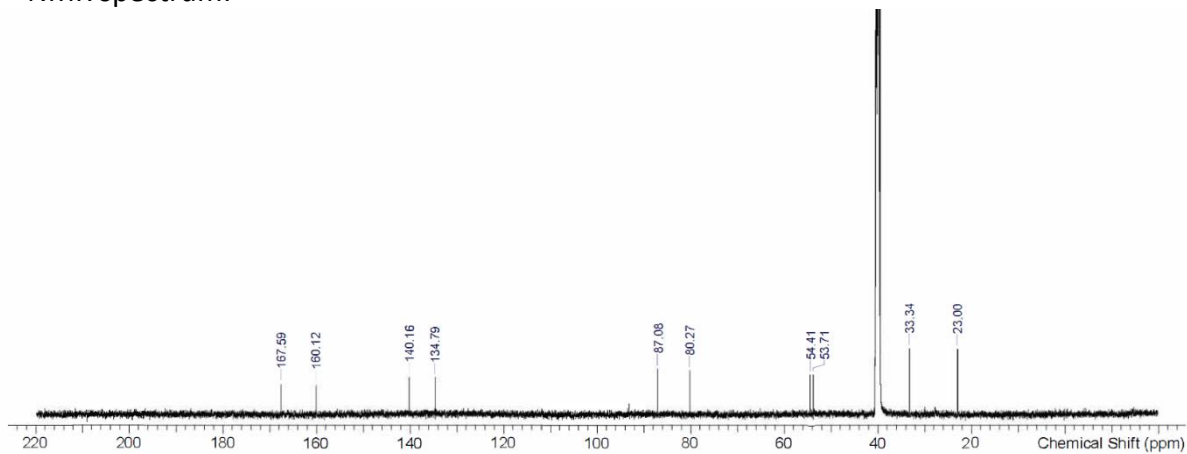
H¹ NMR spectrum:C¹³ NMR spectrum:

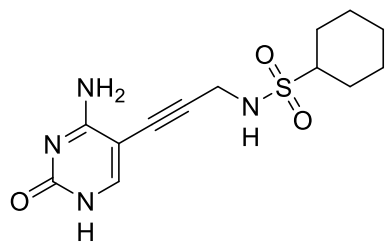


^1H NMR spectrum:

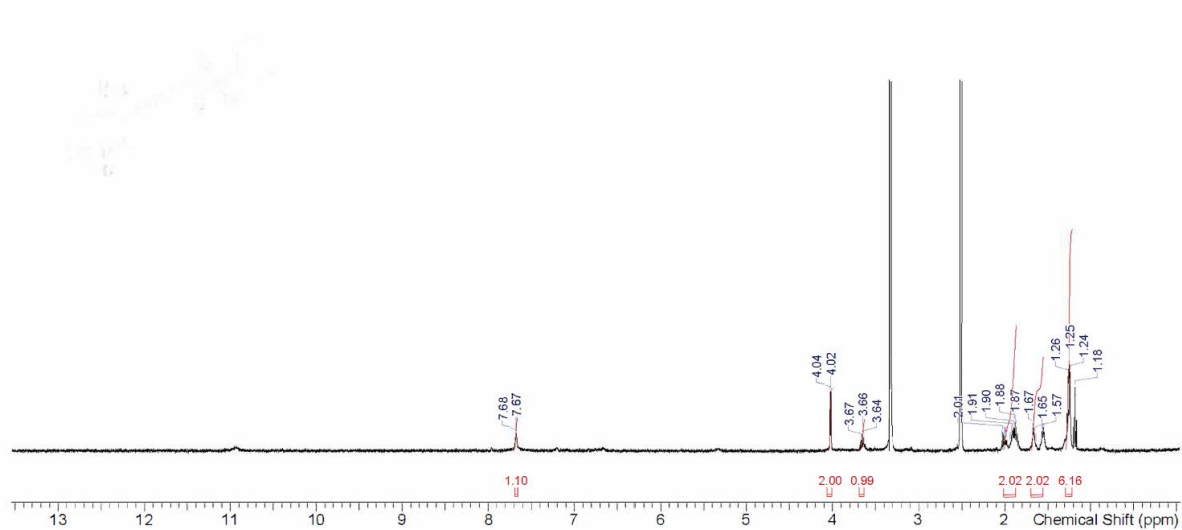


^{13}C NMR spectrum:

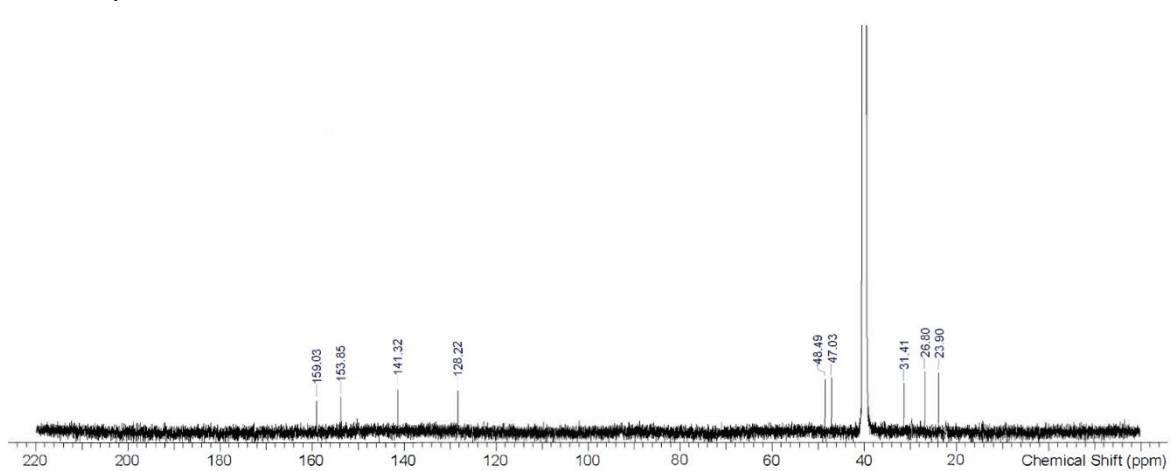


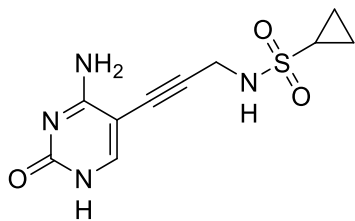


^1H NMR spectrum:

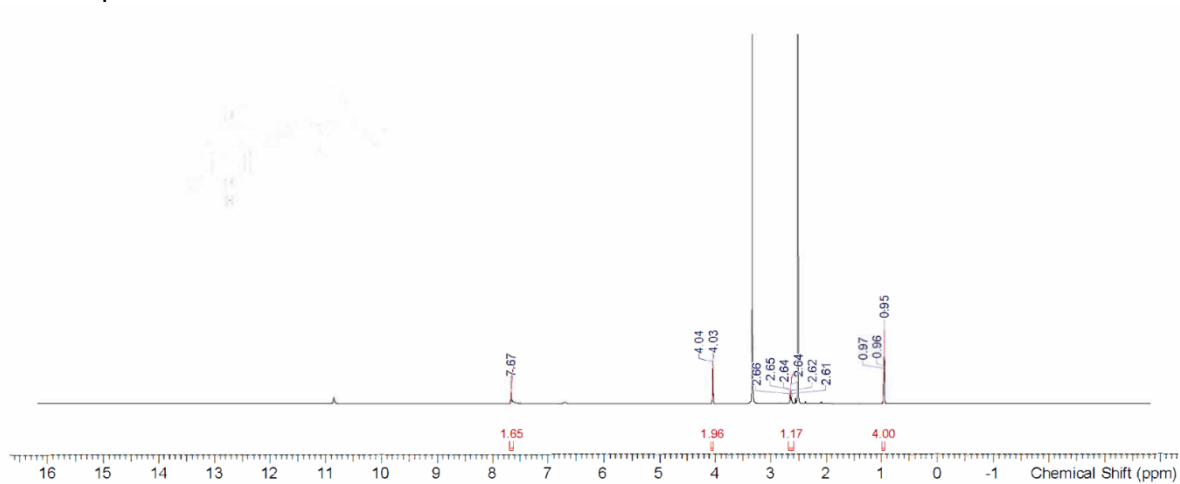


^{13}C NMR spectrum:

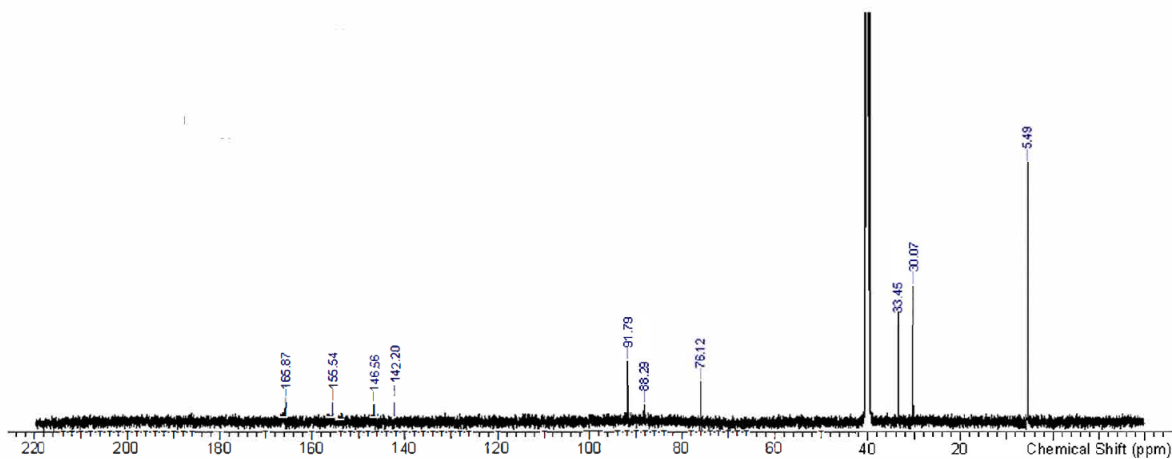


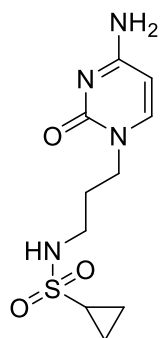


^1H NMR spectrum:

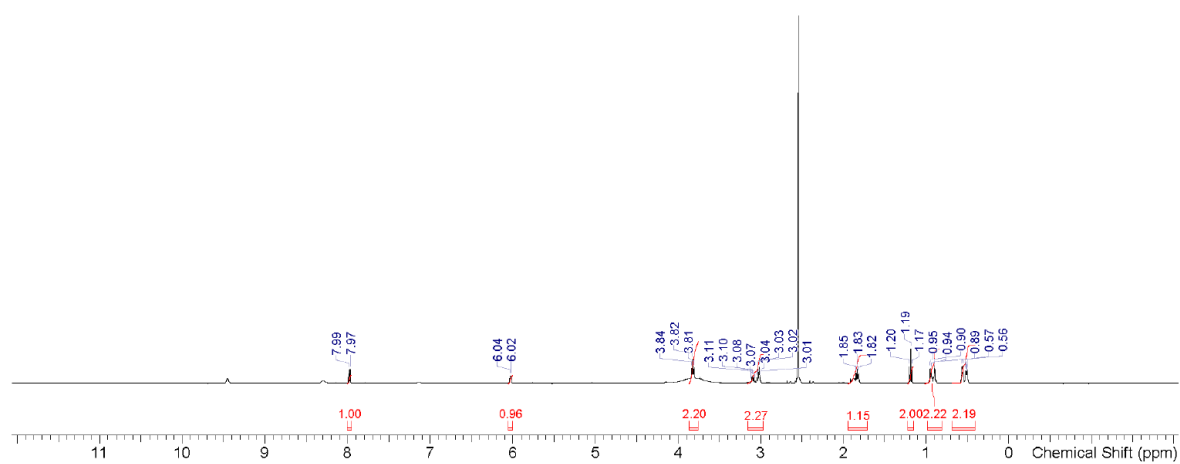


^{13}C NMR spectrum:

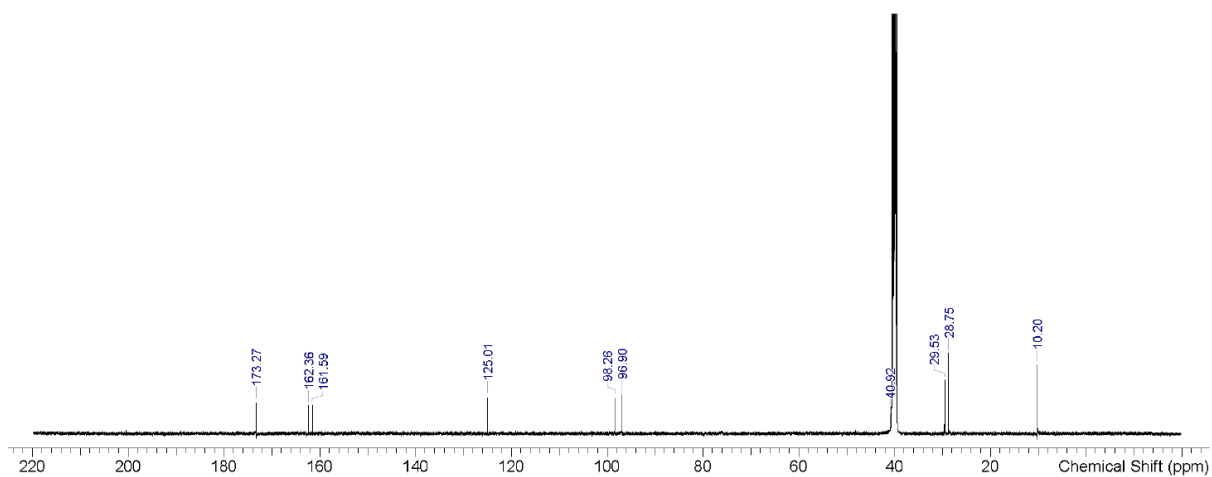


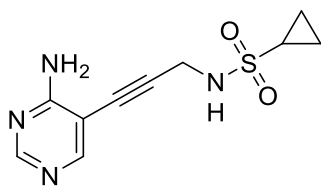


H¹ NMR spectrum:

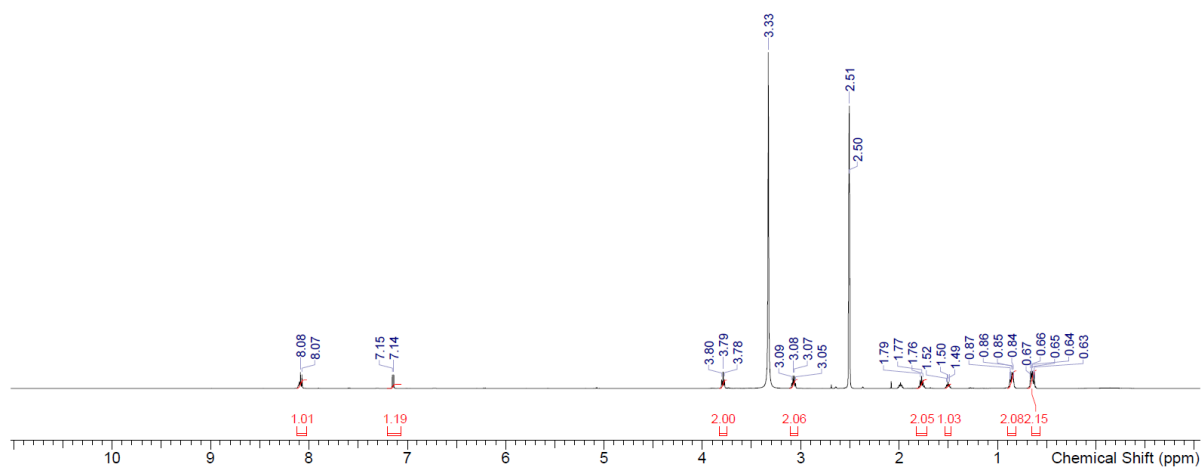


C¹³ NMR spectrum:

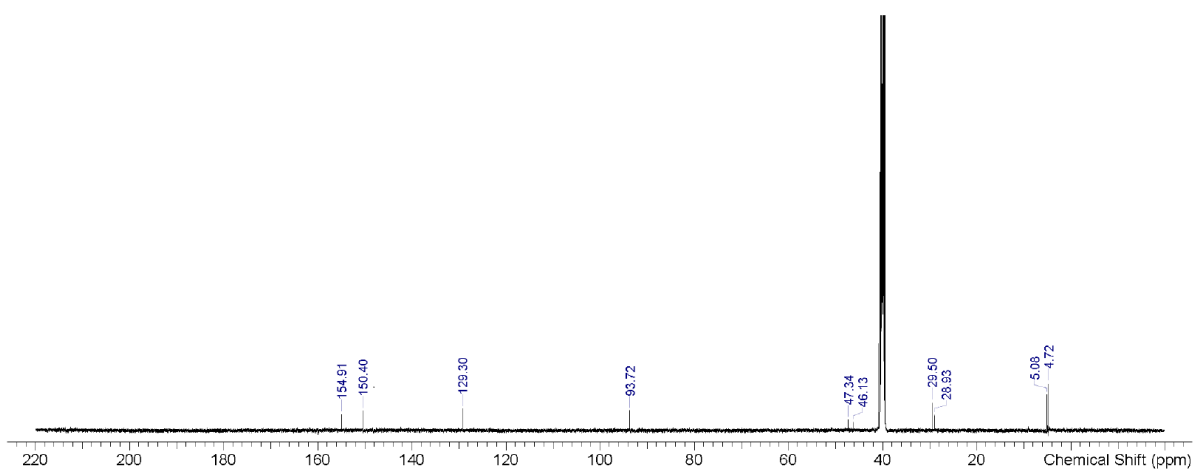


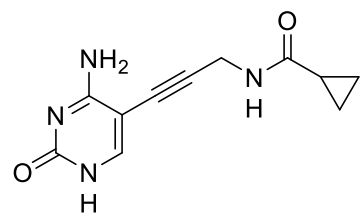


^1H NMR spectrum:

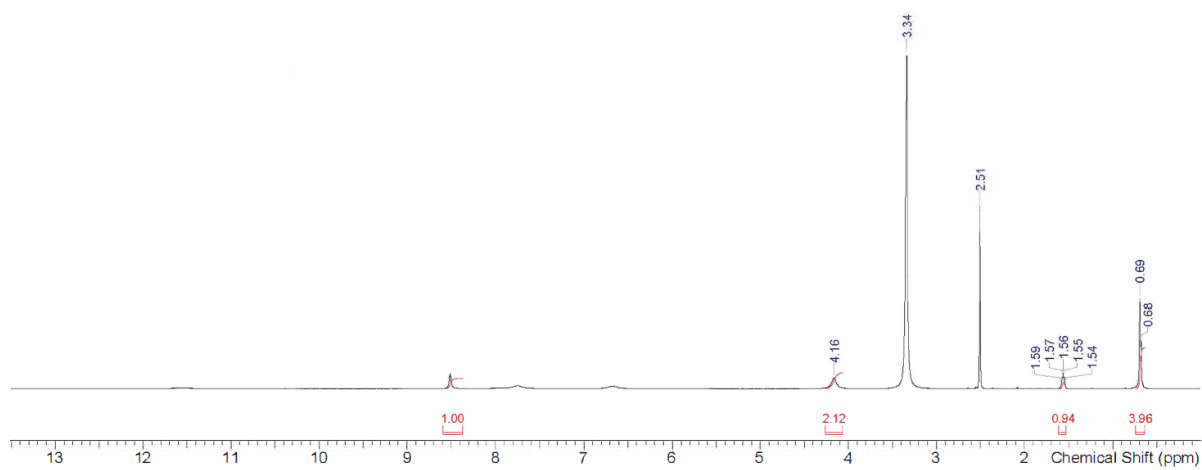


^{13}C NMR spectrum:

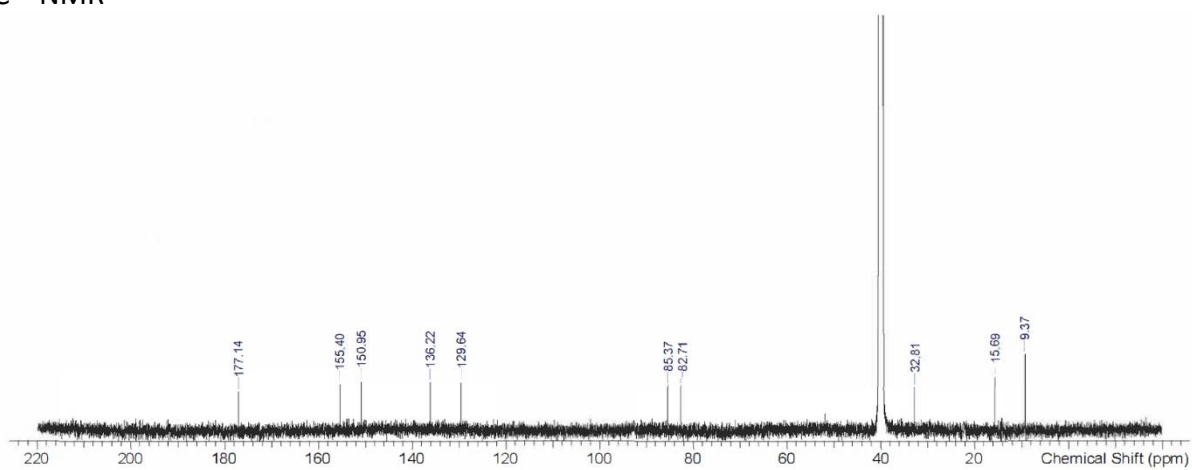


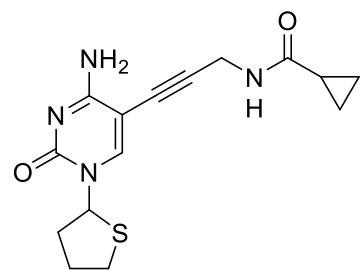


H^1 NMR

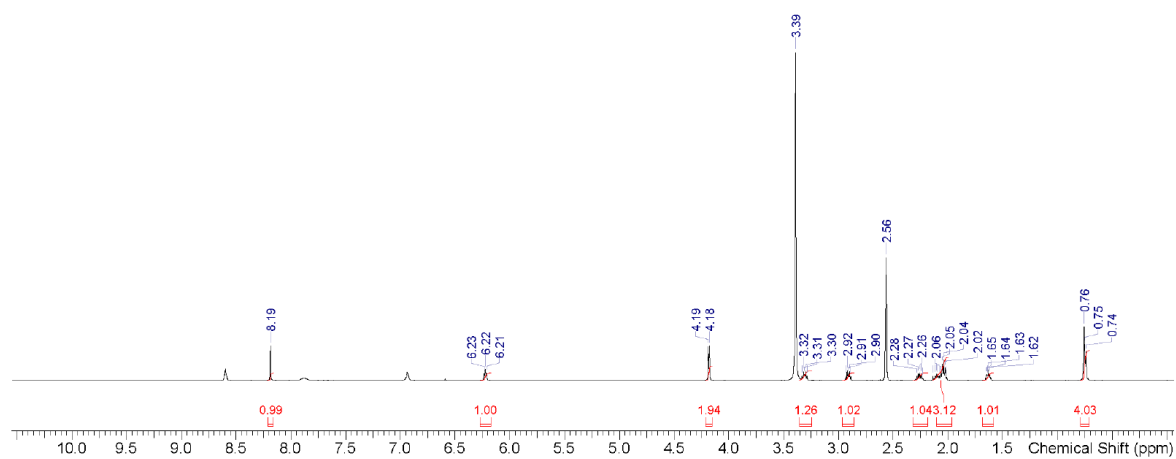


C^{13} NMR

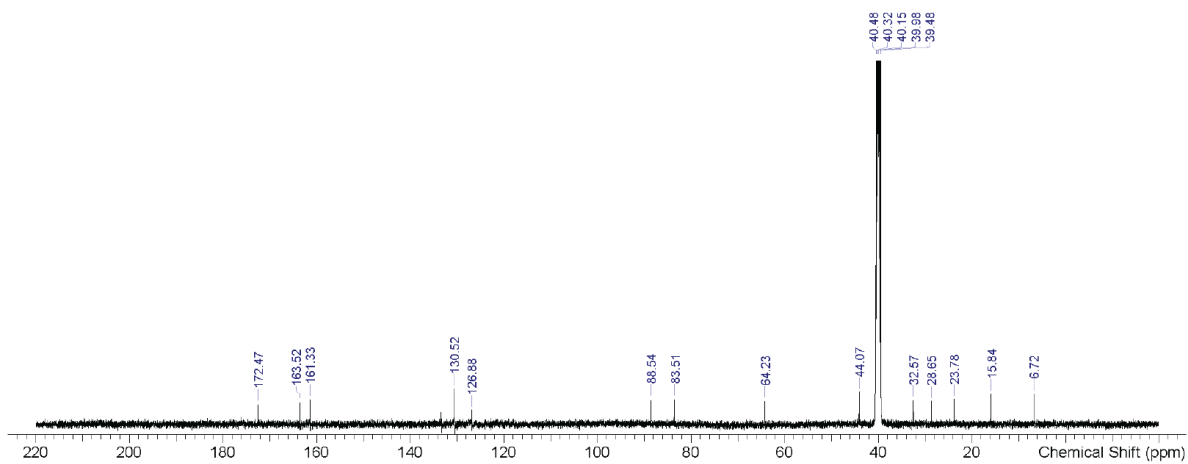




^1H NMR



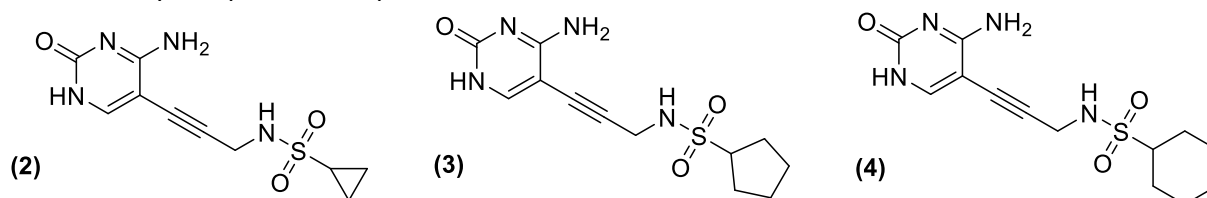
^{13}C NMR



Computational methods: Docking of IspE inhibitors to EclspE and AblspE

All subsequent steps were carried out in MOE v2020.09.

Preparation of ligands. The 2D structures of the three compounds (**2**, **3** & **4**) were sketched using ChemDraw professional 20.0 and were imported into the MOE window. The compounds were subjected to an energy minimization up to a gradient of $0.001 \text{ kcal mol}^{-1} \text{ \AA}^2$ using the MMFF94x force field and R-field solvation model, then they were saved as a .mdb file. The predominant protonation status of the compounds in aqueous medium at pH 7 was calculated via the compute | molecule | wash command in the database viewer window.

**Preparation of protein structures.**EclspE

The X-ray co-crystal structure of *eclspE* with **3** was used for the molecular docking studies. Potential was set up to Amber14:EHT as a force field and R-field for solvation. Missing residues and termini were fixed using the Structure Preparation module in MOE. The Protonate3D module was used to assign and fix hydrogens atoms. Only the receptor atoms were selected in this case as this step would cause the loss of aromaticity in the cytosine ring of the co-crystallized ligand when the ligand was included in the Protonate3D calculation. Tethered energy minimization was then performed via the QuickPrep module, with the Structure Preparation and Protonate3D tick boxes unchecked (as these were already performed manually) while using the default settings for the rest of the module.

AblspE (Uniprot accession code : B2HUM5)

The Swiss-Model structure was downloaded as a .pdb file from the Swiss-Model website via the link provided on the Uniprot web page. Potential was set up to Amber14:EHT as a force field and R-field for solvation. Correction of library errors, and tethered energy minimization of binding site were performed via the QuickPrep module using default settings. The structure was then aligned with the *eclspE* structure in MOE using the sequence viewer's alignment module ("Sequence and Structural" sequence alignment and "Use Current Alignment" superposition methods) with an overall RMSD of 1.118 Å.

Binding site selection.

For *EclspE*, the binding site was defined as the residues within 4.5 Å of the co-crystallized ligand. For the *AblspE* structure, the equivalent residues (as determined by the sequence alignment) were selected as the binding site for docking.

Molecular docking*EclspE*

Molecular docking was performed via the Dock module of MOE. The co-crystallized ligand was selected as a template for docking (similarity mode) while the receptor was set to the prepared *eclspE* structure and the site was selected as described above. Triangle matcher placement was used with 100 poses followed by Rigid Receptor Refinement with 10 poses. The GBVI/WSA dG scoring function was used for both methods with default settings.

AblspE

Molecular docking was performed via the Dock module of MOE. The default docking mode was used while the receptor was set to the prepared *eclspE* structure and the site was selected as described above. Triangle matcher placement was used with 100 poses followed by Rigid Receptor Refinement with 10 poses. The GBVI/WSA dG scoring function was used for both methods with default settings.

Table S 2. Docking scores of the three compounds docked to the three structures of XPO1/CRM1.

Compound	<i>EclspE</i> Docking Score (kcal/mol)	<i>AblspE</i> Docking Score (kcal/mol)
2	-6.95	-6.1
3	-7.08	-6.23
4	-7.04	-6.6

Chapter 3

CONCLUSIONS AND OUTLOOK

3 CONCLUSIONS AND OUTLOOKS

This thesis presented a comprehensive exploration of MEP (2-C-methyl-D-erythritol 4-phosphate) pathway enzymes from selected pathogenic bacteria. The primary focus was on two key enzymes, 1-deoxy-D-xylulose-5-phosphate synthase (DXPS) and 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) homologues, with the aim of deepening our understanding of their structures and functions, as well as their potential as targets for drug development. Main results, conclusions and future perspectives of the thesis are summarized below.

Chapter 2.1. Starting off with DXPS, and to achieve the objective of obtaining structural information from pathogenic organisms targeted by our research group, an optimized DXPS constructs from *P. aeruginosa* was used. This optimization process involved removing flexible regions (loops), with low sequence conservation among DXPS sequences, from the protein sequence, to create a stable construct suitable for crystallization. Comparative kinetic studies between the native enzyme and the modified construct, demonstrated that the truncated construct maintained its enzymatic activity, while improving the stability of the protein.

These sequence optimization (truncation) strategies have allowed the first 3D structure of DXPS from *P. aeruginosa*, as well as another pathogenic microorganism, *K. pneumoniae*, to be obtained, with a resolution of up to 1.8Å. Upon studying their crystal structures, structural differences were observed between the enzyme bound to the cofactor thiamine diphosphate (ThDP) and the *apo* enzyme. This binding induces conformational changes, transitioning the protein between open and closed loop conformations near its active site.

The obtained structural information underscores the critical role of ThDP in stabilizing regions in the active site, including flexible loops that are involved in coordinating the diphosphate moiety. Targeting specific residues in this loop could potentially prevent ThDP from binding by interfering with loop rearrangement, thus preventing the enzyme from adopting its active conformation. These structural insights are important for understanding how DXPS operates and how it might be inhibited.

This newly discovered information about the conformational changes resulting from ThDP binding, in addition to the previously described conformational changes in a different region upon binding of substrates (pyruvate and GAP), showcases the challenges in targeting DXPS using traditional structural-based drug design methods, as the enzyme exists in distinct conformations during catalysis as well as when binding or releasing the cofactor.

Chapter 2.2. This project was part of an effort aimed at designing thiamine analogues to inhibit ThDP-dependent enzymes, including DXPS. This was achieved by exploring various functional groups at the C2 position of the thiazolium, thus targeting the substrate entry site. Additionally, derivatives of ThDP analogues with different central aromatic rings were synthesized, and their inhibitory activity against multiple ThDP-dependent enzymes was

compared. This study established a clear link between the attributes of the central ring and the inhibitory activities of these ThDP competitive inhibitors. The activity of these analogues increased with an increase in the hydrophobicity of the central ring, which occupies a hydrophobic pocket in DXPS. Furthermore, it was demonstrated that expanding through the C2 position of the central ring enhances the potency and selectivity of these analogues against DXPS.

To validate the binding mode of these compounds, a co-crystal structure of DXPS from *K. pneumoniae* with one of these compounds, the thiamine monophosphate (ThMP) analogue, 2-(4-((4-amino-2-methylpyrimidin-5-yl)methyl)-3-methyl-5-propionylthiophen-2-yl) ethyl dihydrogen phosphate, was obtained. The inhibitor binds within the ThDP pocket in a V-shaped conformation, indicating its competitive inhibition of ThDP. Additionally, the 5-acyl group of the ThMP analogue occupies the substrate pocket in a manner similar to the acetyl group in the acetyl ThDP-bound structure, and the inhibitor's monophosphate forms interactions with a Mg²⁺ cation. It is important to note that DXPS enzymes differ significantly in terms of their substrate binding and catalytic action. Therefore, targeting the substrate-binding pocket might result in selective inhibition against DXPS over other ThDP-dependent enzymes.

Chapter 2.3, is a concise report that explores the potential of targeting allosteric binding sites to inhibit DXPS. In this study, a new class of DXPS inhibitors was developed, demonstrating a competitive mode of inhibition with ThDP. An unexpected discovery was made when a co-crystal structure of one of these inhibitors with DXPS from *P. aeruginosa*, at a resolution of 1.9Å, revealed that the inhibitor binds to an allosteric site responsible for shaping the active site, rather than the ThDP-binding site.

As a result, the nature of inhibition with these compounds stems from preventing ThDP from binding by hindering the closure of the diphosphate loop, leading to exposed binding site that is less likely to bind as tightly to ThDP as in its natural state. This discovery opens up new possibilities in designing inhibitors against DXPS, especially considering the high affinity of DXPS to ThDP, which makes it challenging to design small molecules capable of competing within the ThDP binding site.

Nevertheless and to confirm the significance of this allosteric binding site and its potential as a drug target, validation experiments, including site-directed mutagenesis studies, are planned. For example, replacing hydrophilic residues like Ser203 with alanine may disrupt critical hydrogen bonding interactions with the chlorine atom in the first chlorobenzene ring, potentially reducing the activity of these inhibitors and thereby validating their binding site.

Chapter 2.4. Shifting the focus to another enzyme within the same MEP pathway, IspE kinase, IspE from three pathogenic organisms - *E. coli*, *K. pneumoniae*, and *A. baumannii* - were obtained to allow for a comparison of their kinetics, structural characteristics, and inhibition patterns. The crystal structure of IspE from *K. pneumoniae* exhibited a remarkable degree of

similarity to that of *E. coli* IspE. In addition to the structural characteristics, the two homologues also showed similar kinetics and a matching inhibition pattern when tested against different compounds. On the other hand, IspE from *A. baumannii*, aside from showing low sequence homology, demonstrated significant kinetic and structural differences compared to the IspE expressed by both *E. coli* and *K. pneumoniae*. An additional key finding is the different inhibition pattern of *A. baumannii* IspE inhibition influenced by subtle structural differences compared to IspE from *E.coli* and *K. pneumoniae*.

The binding mode of these novel inhibitors were studied in depth by obtained three co-crystal structures with *E.coli* IspE, which was the only IspE homologue in this study that showed crystal packing that allowed for co-crystallization and/or soaking of compounds. The crystals were resolved to a resolution of up to 1.5Å.

After studying the structure-activity relationship (SAR) and the binding of these inhibitors, tailored inhibitor design strategies based on these structural distinctions, which can lead to the development of selective inhibitors for IspE kinase homologues, were proposed. In the case of *A. baumannii* IspE, designing inhibitors requires a different approach due to its unique structural features. In particular, there is a smaller hydrophobic subpocket that affects how compounds bind. For *E. coli* and *K. pneumoniae* IspE, the best approach is to focus on the optimum filling of this hydrophobic pocket, in addition to covalently targeting a cysteine close to this pocket. However, in the case of *A. baumannii* IspE, due to the smaller size of this pocket and its limited capacity to tolerate larger substituents, the focus should shift to the opposite side of the pocket. Here, a glutamine (present in *A. baumannii* IspE) replaces a proline (found in *E. coli* and *K. pneumoniae*). This substitution is noteworthy because glutamine provides an additional carbonyl group that can be targeted for hydrogen bonding. While, proline plays a crucial role in how the active sites of *E. coli* and *K. pneumoniae* IspE interact with the ribose component in the substrate CDP-ME, as well as ligands designed to target this region of the substrate pocket.

In summary, this thesis represents a contribution to our understanding of MEP pathway enzymes and their potential as targets for drug development. It introduces novel concepts, such as DXPS allosteric binding site targeting, which could lead to innovative approaches in drug discovery targeting this class of enzymes. Moreover, it underscores the importance of considering structural variations among enzyme homologues when designing inhibitors, ultimately advancing our capabilities in fighting drug-resistant pathogens.