

1-Deoxy-D-xylulose-5-phosphate synthase: a promising target for novel antibiotics and antimalarials

Author: Fabian Mulder
Supervisor: Rick Oerlemans
Date: 22-2-2019
Department of Drug Design
University of Groningen
Groningen

Table of contents

Abstract	2
Introduction.....	2
Possible important residues for substrate binding	7
Structural differences between ecDXS and drDXS	9
Sequence analysis for multiple forms of DXS.....	13
Structural modelling <i>P. falciparum</i> DXS	14
Structural modelling <i>M. tuberculosis</i> DXS.....	15
Alkylacetylphosphonates	15
Different applications for DXS targeting compounds	16
Increasing artemisinin production by overexpression of <i>Artemisia annua</i> DXS.....	17
Conclusion	17
References.....	18

Abstract

The enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXS) catalyses the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) through the formation of a ternary complex consisting of DXS, D-glyceraldehyde 3-phosphate and a complex of pyruvate and thiamine diphosphate. This formation of DXP is the initial step in the methylerythritol phosphate (MEP) pathway. The MEP pathway is a biosynthetic route used by many organisms, such as *Plasmodium falciparum*, known to cause malaria, *Mycobacterium tuberculosis*, known to cause tuberculosis, and *Deinococcus radiodurans*, for the production of isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP). IDP and DMADP are used for the production of terpenoids, which fulfil a wide range of functions in the metabolism and survival of the organism. Humans utilize the mevalonate pathway instead of the MEP pathway and thus selective inhibition of the MEP pathway through targeting of DXS is an interesting mode of action for new antibiotics and antimalarials. Protein sequence- and crystal structure analyses are required for the design of compounds targeting DXS and important residues involved with the substrate binding of DXS might be promising targets. There are two crystal structures available of DXS, from *E. coli* and *D. radiodurans*. Designing novel antibiotic compounds is necessary to combat antibiotic resistance in pathogenic bacteria and targeting DXS could provide a novel avenue in the development of new antibiotics. *P. falciparum* also utilizes the MEP pathway, but its DXS is noticeably different compared to other DXS types shown in this study. Even though there is no crystal structure of *P. falciparum* available, structural modelling analyses provide some insight into the characteristics of this DXS type.

Introduction

The enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXS) catalyses the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) which is an initial step and the rate limiting step in the methylerythritol phosphate (MEP) pathway. (Kudoh, Hotta, Sekine, Fujii, Uchida, Kubota, ... , & Ihara, 2017; Zhang, Liu, Xia, Zeng, Xiang, Zhu, ... , & Liao, 2018) This is a crucial mechanism for a wide range of non-mammalian, in some cases pathogenic, organisms in the production of isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP) (Bartee, & Freel Meyers, 2018; Banerjee, Preiser, & Sharkey, 2016). IDP and DMADP are building blocks containing five carbon atoms and they are used for the production of terpenoids, a wide range of molecules with many important functions in the cells of these organisms, such as vitamin E, cholesterol, hormones, carotenoids, pigments and chlorophyll (Ramamoorthy, Handa, Merkler, & Guida, 2014; Battistini, Shoji, Handa, Breydo, & Merkler, 2016). Humans utilize the Mevalonate pathway instead of the MEP pathway, but some organisms that do utilize the MEP pathway in the production of terpenoids are *Mycobacterium tuberculosis*, *Plasmodium falciparum*, *Deinococcus radiodurans* and *Escherichia coli*. (Ramamoorthy, Handa, Merkler, & Guida, 2014) DXP is produced with CO₂ as a byproduct through a thiamine diphosphate- (ThDP) and Mg(II) or Mn(II) dependent decarboxylation (Handa, Dempsey, Ramamoorthy, Cook, Guida, Spradling, & Merkler, 2018; Goswami, 2017) and condensation of D-glyceraldehyde 3-phosphate (GAP) and pyruvate and DXP is also a precursor in the biosynthesis of vitamin B₆ and vitamin B₁ (Basta, Patel, Kakalis, Jordan, & Meyers, 2014; Xiang, Usunow, Lange, Busch, & Tong, 2007; Goswami, 2017).

The MEP pathway consists of seven enzymatically catalysed reactions (figure 1). 1) The formation of DXP from pyruvate and GAP by DXS. 2) The conversion of DXP into MEP by 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR). 3) The conversion of MEP into methylerythritol 2,4-cyclodiphosphate (MEcDP) through cytidylation by 4-diphosphocytidyl-2-C-methylerythritol synthase/2-C-methyl-D-erythritol 4-phosphate (CMS/MCT), 4) phosphorylation by 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (CMK) and 5) cyclization by 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS). 6) The conversion of MEcDP into hydroxymethylbutenyl diphosphate (HMBDP) by HMBDP synthase (HDS). 7) The reduction of HMBDP to IDP and DMADP by HMBDP reductase (HDR). The IDP and DMADP products can also be isomerized by isopentyl diphosphate isomerase (IDI). (Banerjee, & Sharkey, 2014; Zhao, Chang, Xiao, Liu, & Liu, 2013; Frank, & Groll, 2016; Matsue, Mizuno, Tomita, Asami, Nishiyama, & Kuzuyama, 2010)

The first step of the formation of DXP is the formation of a ThDP-pyruvate complex, after which the COO⁻ group of the bound pyruvate is released as CO₂ (Bartee, & Freil Meyers, 2018; Zhao, Chang, Xiao, Liu, & Liu, 2013; Frank, & Groll, 2016). Subsequently GAP binds to the complex, and in the final step ThDP is released and DXP has been formed (figure 2). This process is unusual, as a ternary complex is formed between DXS, the ThDP-pyruvate complex, C2α-lactylthiamin diphosphate (LThDP), and GAP, while other observed ThDP dependent pyruvate decarboxylase enzymes function through the classical ordered mechanism or through the so-called ping-pong mechanism where the binding of substrate and acceptor only occurs after release of CO₂ (Bartee, & Freil Meyers, 2018; Hailes, Rother, Müller, Westphal, Ward, Pleiss, ... , & Pohl, 2013; Frank, Leeper, & Luisi, 2007).

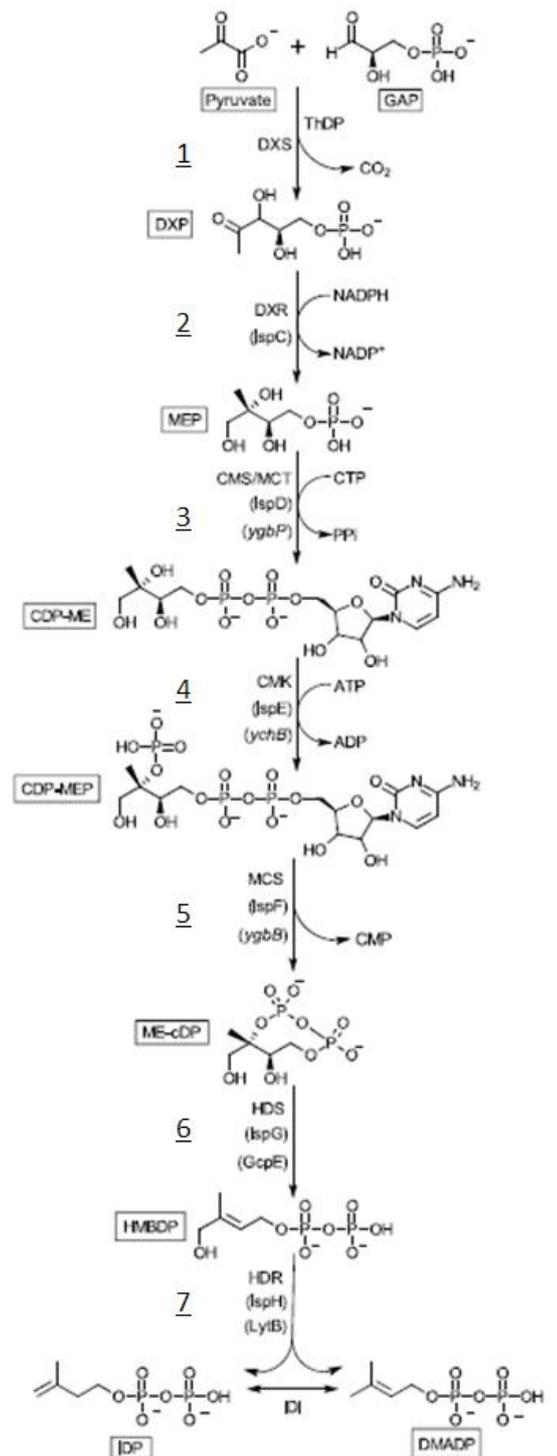


Figure 1: The MEP pathway with important enzymes and metabolites. DXR = 1-deoxy-D-xylulose 5-phosphate reductoisomerase, CMS/MCT = 4-diphosphocytidyl-2-C-methylerythritol synthase/2-C-methyl-D-erythritol 4-phosphate cytidyltransferase, CDP-ME = diphosphocytidyl methylerythritol, CMK = 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase, CDP-MEP = diphosphocytidyl methylerythritol 2-phosphate, MCS = 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, ME-cDP = methylerythritol 2,4-cyclodiphosphate, HDS = 4-hydroxy-3-methylbut-2-enyl diphosphate synthase, HMBDP = hydroxymethylbutenyl diphosphate, HDR = 4-hydroxy-3-methylbut-2-enyl diphosphate reductase. (Banerjee, & Sharkey, 2014)

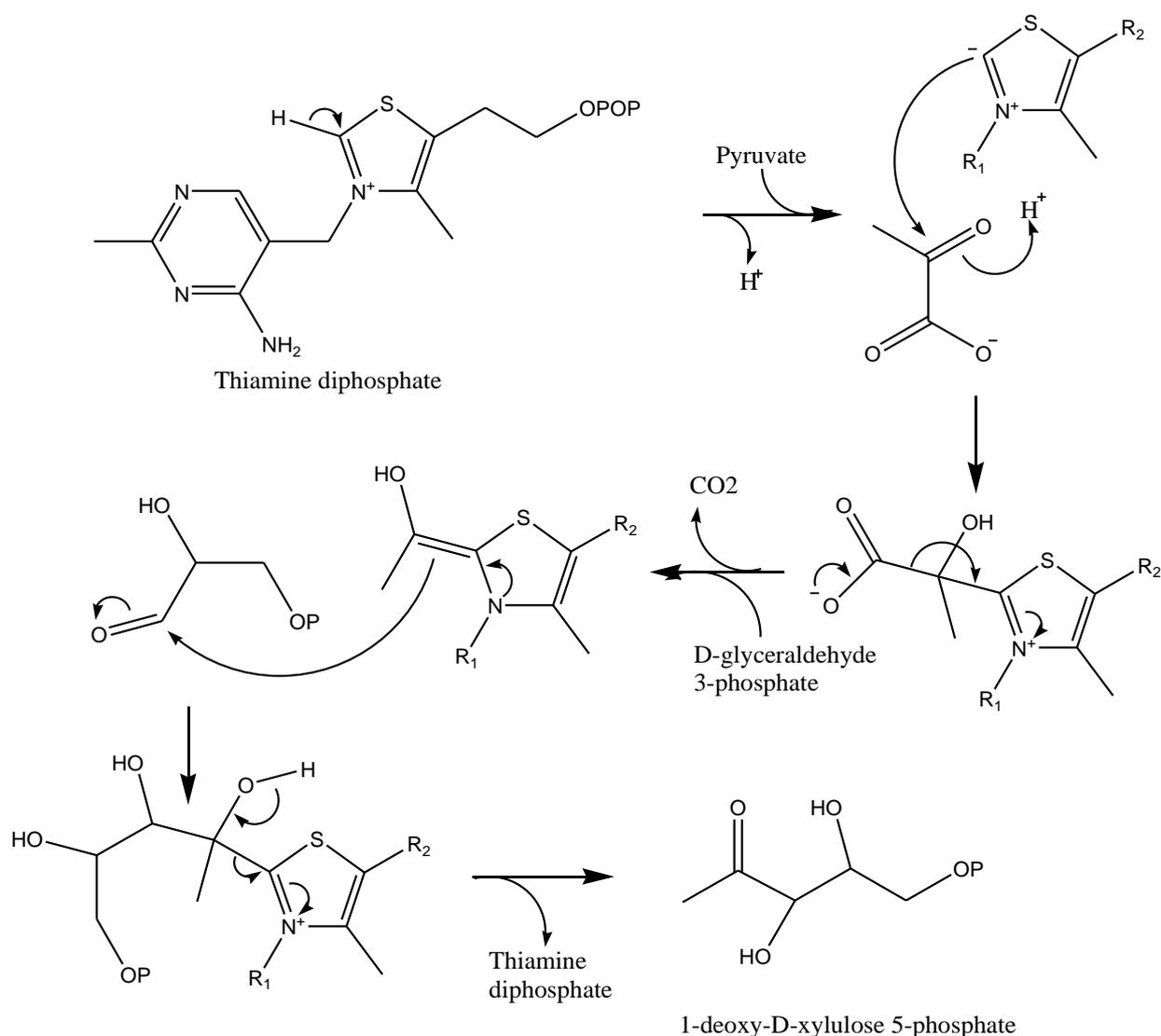


Figure 2: Reaction scheme for the formation of DXP. Reaction scheme was drawn using ChemDraw Ultra 7.0.1 (Developed by Cambridge Soft Corp.)

The necessity of GAP binding is observed by a significant increase of the decarboxylation rate of pyruvate when GAP is bound, in comparison to the decarboxylation rate in the absence of GAP (Figure 3) (Bartee, & Freel Meyers, 2018). It has been shown that DXS catalyses the decarboxylation of pyruvate in conditions in which LThDP accumulates. (DeColli, Nemeria, Majumdar, Gerfen, Jordan, & Meyers, 2018) The need to form the ternary complex is consistent with the observed large volume of the DXS active site as compared to transketolase- (TK) and pyruvate dehydrogenase (PDH) active sites. (Bartee, & Freel Meyers, 2018; Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017)

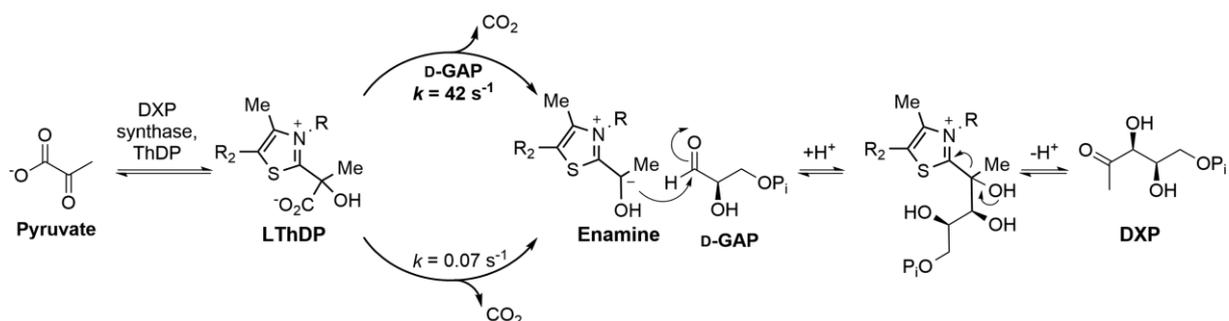


Figure 3 Mechanism of DXS with the decarboxylation rates with- and without GAP, 42s^{-1} and $0,07\text{s}^{-1}$ respectively. (Bartee, & Freel Meyers, 2018)

After DXP is formed, it can be used for three different biosynthetic pathways: the biosynthesis of terpenoids, the biosynthesis of ThDP and the biosynthesis of pyridoxal phosphate (PLP or vitamin B₆) (figure 4). (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017)

The importance of DXS for the MEP pathway in pathogenic organisms and the absence of the MEP pathway in mammals cause it to be an interesting target for novel antibiotics, antimalarials and herbicides (Banerjee, & Sharkey, 2014). Knowledge of the crystal structure and the active site of DXS is important for the design of DXS targeting drugs and important residues in the active centre could be used as targets for novel compounds. The crystal structures of DXS are available from *D. radiodurans* (drDXS) and *E. coli* (ecDXS), which are both homodimers. Possible important residues in ecDXS are Arg420, Arg478 and Tyr392.

(Brammer, Smith, Wade, & Meyers, 2011; Xiang, Usunow, Lange, Busch, & Tong, 2007; Basta, Patel, Kakalis, Jordan, & Meyers, 2014) *Populus trichocarpa* DXS (ptDXS) was found to be inhibited by IDP and DMADP, while the metabolites compete with ThDP for binding with ptDXS and it was indicated that ThDP dependent enzymes might often be affected by DMADP and IDP. (Banerjee, Wu, Banerjee, Yan, & Sharkey, 2013)

Ketoclozazole, a herbicide, is shown to inhibit *Chlamydomonas* DXS (Hayashi, Kato, Kuzuyama, Sato, & Ohkanda, 2013; Mueller, Schwender, Zeidler, & Lichtenthaler, 2000) and was found to suppress the growth of the pathogenic bacterium, *Haemophilus influenzae*. (Matsue, Mizuno, Tomita, Asami, Nishiyama, & Kuzuyama, 2010) The inhibition kinetics of ketoclozazole suggested that it binds to an unidentified binding site, which is different from the pyruvate binding site and from the GAP binding site on DXS. Therefore the inhibition of DXS by ketoclozazole is uncompetitive. It is also found that ketoclozazole can only bind to the enzyme-substrate complex and not to free DXS. (Matsue, Mizuno, Tomita, Asami, Nishiyama, & Kuzuyama, 2010) A possible difference in sensitivity to ketoclozazole has also been found between *H. influenzae* and *E. coli* which could indicate that the ketoclozazole binding site could involve residues which are not conserved between these two organisms. (Matsue, Mizuno, Tomita, Asami, Nishiyama, & Kuzuyama, 2010)

A possible way to target DXS is by using alkylacetylphosphonates (alkylAPs) (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017), that act as a substrate by mimicking pyruvate and forming a covalent bond with ThDP which results in a phosphonolactyl ThDP (PThDP) dead-end intermediate (Figure 5) (Bartee, & Freel Meyers, 2018). Another way to target DXS might be through a compound that has properties which act as donor and acceptor substrate simultaneously, which

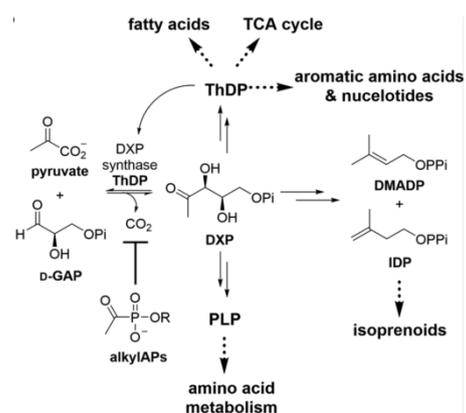
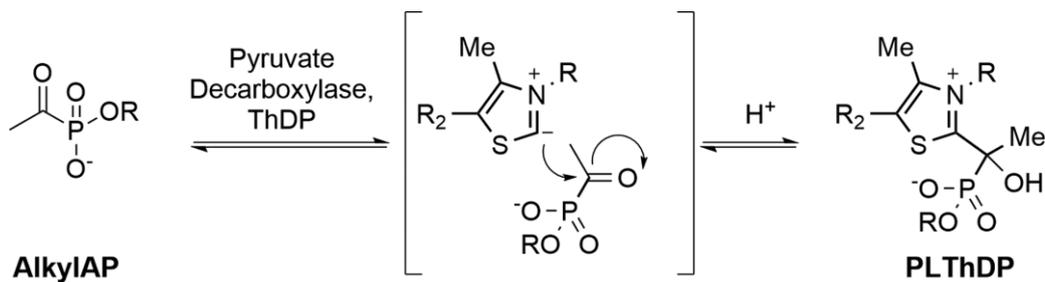


Figure 4: The three different biosynthetic pathways DXP can be used for (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017)

could result in a highly potent and selective compound for DXS inhibition. (Bartee, & Freel Meyers, 2018)



AlkylIAP
PLThDP
 Figure 5: Mechanism in which alkylIAP mimics pyruvate and forms a covalent dead-end intermediate, PLThDP. (Bartee, & Freel Meyers, 2018)

There are two crystal structures available for DXS from *Escherichia coli* and *Deinococcus radiodurans* with Protein Data Bank (PDB) codes 2O1S and 2O1X respectively. The drDXS crystal structure is complete, but the ecDXS crystal structure is incomplete due to the occurrence of in situ proteolysis during its crystallisation by a fungal protease. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015; Xiang, Usunow, Lange, Busch, & Tong, 2007)

It has been found that the *Arabidopsis thaliana* J-protein J20 interacts with inactive forms of DXS and delivers them to the Hsp70 chaperone where it will be activated, by folding or refolding, or degraded, which involves unfolding. (Pulido, Toledo-Ortiz, Phillips, Wright, & Rodríguez-Concepción, 2013) Hsp70 acts together with specific J-proteins and Hsp100 chaperones to recognise and deliver DXS to degradation, through ClpC1, or reactivation, through ClpB3. (Pulido, Llamas, Llorente, Ventura, Wright, & Rodríguez-Concepción, 2016) The CLA1 gene, encodes for DXS from *Arabidopsis thaliana* (Estévez, Cantero, Romero, Kawaide, Jiménez, Kuzuyama, ... , & León, 2000) and is located on position 56 of chromosome IV in *Arabidopsis thaliana*. (Crowell, Packard, Pierson, Giner, Downes, & Chary, 2003) Insight into the mechanisms involved in reactivation of DXS and degradation of DXS might present new ways of inhibiting DXS through these mechanism, by inhibiting the reactivation and/or inducing the degradation of DXS. These indirect DXS inhibitions could therefore form new ways of targeting DXS and these different targeting mechanisms could be important to combat possible future pathogens that will be resistant to direct DXS targeting compounds.

Possible important residues for substrate binding

Arg420 and Arg478 in ecDXS, corresponding to Arg423 and Arg480 in *D. radiodurans* (figure 6) have been proposed to interact with the phosphoryl group in GAP. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014; Xiang, Usunow, Lange, Busch, & Tong, 2007) It was suggested that Arg478 and Arg420 might be important in the mechanism of GAP binding and the catalysis. (Xiang, Usunow, Lange, Busch, & Tong, 2007) To analyse this possibility R478A and R420A ecDXS productions and kinetic characterisations have been performed (table 1). This resulted in a significant decrease in K_m^{GAP} for R478A in comparison with the wild type DXS (table 1), suggesting that Arg478 plays a role in GAP binding. The analysis also showed a similar

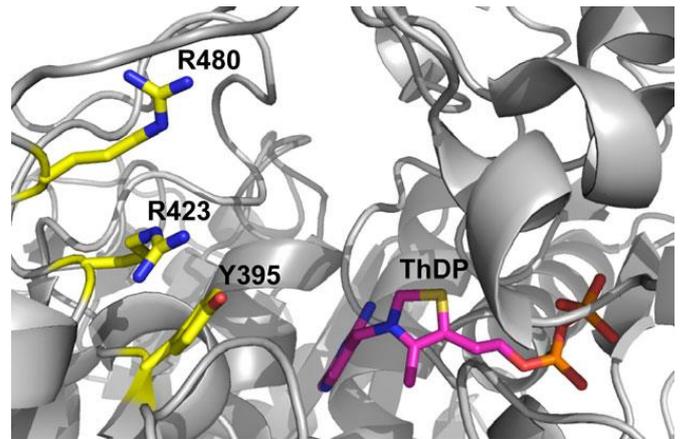


Figure 6: Active site residues R423, R480 and Y395 in *D. radiodurans* (corresponding to R420, R478 and Y392 in *E. coli*, respectively) with bound ThDP (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

pyruvate affinity (K_m) and substrate turnover (K_{cat}) for R478A as was found in the wild type DXS, which suggests that Arg478 is not important for catalysis. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) It was also found that the LThDP decarboxylation rate for the R478A DXS was similar to that of the wild type DXS. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) The LThDP formation for the R420A enzyme was similar to that of the wild type enzyme, but in the absence of GAP the decarboxylation rate for LThDP could not be measured within a 50 seconds time frame (table 2) which suggests that the R420A residue might have a significant stabilizing effect on GAP. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) The LThDP decarboxylation rate for R420A in the presence of GAP was also measured and was found to be similar to that of the wild type DXS (table 2), suggesting that the residue is not important for this catalysis. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) The significantly lower specificity constants, K_{cat}/K_m^{GAP} , supported the expectation that Arg420 and Arg478 would have an effect on the GAP binding. It is suggested that GAP fulfils two functions in the DXS reaction mechanism: increasing the decarboxylation process of LThDP by the DXS-LThDP-GAP complex and acting as the acceptor in the carbonylation step by the DXS-enamine-GAP complex. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) Because of these two functions, there are also two K_m -values which should be distinguished. The K_m^{GAP} was elevated in both R420A DXS and R478A DXS (table 1) which would be expected to result in increased GAP-promoted LThDP decarboxylation rates. However, the measured decarboxylation rates were found to be similar to that of the wild type DXS (table 2). This finding contributes to the suggestion that GAP fulfils two functions and in this case the elevated K_m^{GAP} is caused by a higher affinity of GAP in the carbonylation step while the affinity of GAP in the decarboxylation step remains unaltered. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

Table 1: Summary of the wild type DXS and the altered DXS enzymes. WT, wild type; ND, not defined. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

Enzyme	K_m^{GraP} (μM)	K_d^{GraP} (μM)	$K_m^{pyruvate}$ (μM)	$K_d^{pyruvate}$ (μM)	k_{cat} (min^{-1})	k_{cat}/K_m^{GraP} ($\mu M^{-1} min^{-1}$)	$k_{cat}/K_m^{pyruvate}$ ($\mu M^{-1} min^{-1}$)
WT	23.5 ± 1.7^a	2000 ± 500^b	49 ± 8^a	230 ± 70^b	154 ± 7^a	6.8 ± 0.4^a	3.2 ± 0.4^a
R478A	1600 ± 300	ND ^c	54 ± 7	80 ± 30^d	200 ± 10	0.14 ± 0.02	3.1 ± 0.5
R420A	ND ^c	ND ^c	ND ^c	37 ± 8^d	ND ^c	0.0018 ± 0.0003^e	ND ^c
Y392F	210 ± 20	ND	73 ± 4	230 ± 60^b	179 ± 4	0.91 ± 0.06	2.37 ± 0.15

It has been stated that DXS utilises a random sequential mechanism through the formation of a ternary complex of DXS-LThDP-GAP (Bartee, & Freil Meyers, 2018; Battistini, Shoji, Handa, Breydo, & Merkle, 2016) and this mechanism has already been proposed in previous research. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) The suggested ternary complex formation caused a focus on possible residues which might play a role in pyruvate binding and ThDP-bound intermediate stabilization. For suggestions of such possible residues, the crystal structure of a complex of the E1 component of *E. coli* PDH (PDHc E1) with C2 α -phosphono-LThDP was examined. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) This complex suggests a stabilisation and positioning of this intermediate through the formation of a hydrogen bond between an oxygen atom in one of the phosphoryl groups and the Tyr599 residue in PDHc E1. This Tyr599 residue was predicted to be an analogue to Tyr392 in ecDXS (Xiang, Usunow, Lange, Busch, & Tong, 2007) and therefore Y392F ecDXS production and kinetic characterisation has been performed (table 1). The substrate turnover was found to be similar to that of the wild type DXS. The small increase in K_m^{pyruvate} compared to the wild type DXS suggested that the hydroxyl groups of Tyr392 is not essential for but might contribute to the binding of ThDP-bound intermediates or pyruvate. This is also suggested by the similar affinity of pyruvate binding to Y392F DXS and to the wild type DXS. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) Stopped flow CD analyses pointed out that the LThDP formation of Y392F DXS was similar to that of the wild type DXS, but the LThDP decarboxylation rate in absence of GAP was lower (table 2), suggesting that Y392F might have a stabilising effect on LThDP. LThDP stability on Y392F was also indicated by proton NMR spectrometry. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) The decreased LThDP decarboxylation rates for the R420A DXS and the Y392F DXS in the absence of GAP suggest that these amino acid changes result in a heightened stability of LThDP, mostly for the R420A DXS (table 2). Therefore it has been suggested that Arg420 and Tyr392 may cause a decreased occurrence of decarboxylation in the absence of GAP when the DXS-LThDP-GAP complex cannot be formed. This decreased decarboxylation could be caused by the formation of a barrier by Arg420 and Tyr392, which could be even more so in the cases of R420A DXS and Y392F DXS due to the observed heightened stability of LThDP (table 2). (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) It could be that the presence of GAP, when in the DXS-LThDP-GAP complex, causes the R420A and Y392F residues to be repositioned, resulting in a similar LThDP decarboxylation rate as compared to the wild type DXS. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

Table 2: Microscopic rate constants for LThDP formation, LThDP decarboxylation and DXP formation on the wild type DXS and the altered DXS enzymes (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

Enzymes	Formation of LThDP (s ⁻¹)		LThDP decarboxylation (s ⁻¹)		Formation of DXP (s ⁻¹)
	-GraP ^a	+GraP ^b	-GraP ^c	+GraP ^d	
Wild type ^e	$k_1 = 1.39 \pm 0.05$ ($k_1' = 0.16 \pm 0.02$)	$k_1 = 0.68 \pm 0.01$	$k_2 = 0.07 \pm 0.006$	$k_2 = 42.0 \pm 1.0$	$k_4 = 1.24 \pm 0.03$ ($k_4' = 0.29 \pm 0.01$)
R478A	$k_1 = 0.85 \pm 0.03$ ($k_1' = 0.29 \pm 0.02$)	$k_1 = 0.84 \pm 0.004$	$k_2 = 0.04 \pm 0.004$	$k_2 = 36.5 \pm 1.1$	$k_4 = 0.78 \pm 0.08$ ($k_4' = 0.26 \pm 0.02$)
R420A	$k_1 = 1.24 \pm 0.01$ ($k_1' = 0.28 \pm 0.002$)	$k_1 = 0.67 \pm 0.003$	No decomposition	$k_2 = 38.8 \pm 0.9$	$k_4 = 0.71 \pm 0.02$
Y392F	$k_1 = 1.06 \pm 0.03$ ($k_1' = 0.3 \pm 0.01$)	$k_1 = 0.82 \pm 0.01$	$k_2 = 0.01 \pm 0.003$	$k_2 = 39.5 \pm 1.4$	$k_4 = 1.34 \pm 0.12$ ($k_4' = 0.34 \pm 0.01$)

Another residue is ecDXS which has been found to be important is H49. (Querol, Rodríguez-Concepción, Boronat, & Imperial, 2001) It has been found that replacing this histidine with a glutamate, H49Q ecDXS, resulted in a undetected DXS activity. (Querol-Audí, Boronat, Centelles, & Imperial, 2014) It has been supported that the H49 residue in ecDXS and the H30 residue in yeast TK

are conserved equivalent residues and could have similar roles in catalysis which would indicate that the loss of H49Q ecDXS activity could be caused by a disturbed proton transfer during catalysis. (Querol, Rodríguez-Concepción, Boronat, & Imperial, 2001) In yeast TK there is another histidine residue, H263, is thought to be important for the catalysis as it works together with H30, however such a residue is not determined in ecDXS, which could indicate that this second residue is not required for the catalysis in ecDXS. However, the two residues H195 and H232 have been found in ecDXS which could also act as this second residue. (Querol, Rodríguez-Concepción, Boronat, & Imperial, 2001) E370 in ecDXS was found to be essential for catalytic activity, while it is found to be involved in cofactor deprotonation, which is the first step in enzymatic thiamine catalysis. (Querol-Audí, Boronat, Centelles, & Imperial, 2014) H431 in ecDXS was indicated to contribute to transition state stability for ecDXS. It has been proposed that D427 in ecDXS plays a role in substrate binding and stereoisomer selection and it has been observed that D427N DXS and D427A DXS mutants showed a significant DXS activity decrease. (Querol-Audí, Boronat, Centelles, & Imperial, 2014)

Structural differences between ecDXS and drDXS

DXS is found to be a novel class of transketolase-like proteins. It is a homodimer and each monomer can be divided into three domains, which all have an α/β fold where a central β -sheet is situated between α -helices. Domain 1 consists of residues 1-319 and has a five-stranded parallel β -sheet, domain 2 consists of residues 320-495 and has a six-stranded parallel β -sheet and domain 3 has a five-stranded β -sheet of which the first strand is not parallel to the other strands. The dimer interface is mostly hydrophobic and the monomers have a side-by-side arrangement in the dimer and thus each domain is in contact with the same domain of the other monomer. (Xiang, Usunow, Lange, Busch, & Tong, 2007)

Several structural homologues have been found to show structural similarity to DXS domains. Some of these homologues are TK, the E1 subunit of PDH and 2-oxoisovalerate dehydrogenase. It has been found that the difference in organisation of the DXS dimer as opposed to those of TK and PHD, is that the DXS active site is situated within the same monomer and in PDH and TK the active site is situated at the interface of the two monomers. (Xiang, Usunow, Lange, Busch, & Tong, 2007)

To further assess the structural differences between ecDXS and drDXS, the crystal structures were obtained from the PDB, with the identification codes 2O1S and 2O1X for ecDXS and drDXS respectively. These structures have been visualized using Pymol (developed by Schrödinger) (figure 7). It was noticed that drDXS had a loop (residues 300-323) which was not shown in the ecDXS, likely because in ecDXS this loop could not be ordered properly in the crystal structure due to variability in the positioning of the loop and therefore is not shown in Pymol.

In figure 8 the residues were examined more closely and it was noticed that the residues 300-323 loop of drDXS is situated in front of the red and magenta coloured residues, which could possibly affect compounds targeting one of these residues (Arg420, Arg478 and Tyr392 for ecDXS) as it might cause steric hinder for a such a compound. Also noticeable is the difference in orientation of the Arg478 (ecDXS) and Arg480 (drDXS). This difference in orientation shown clearer in figure 9 and here it is also shown that the other two aligned residue couples, Arg420 (ecDXS) with Arg423 (drDXS) and Tyr392 (ecDXS) with Tyr395 (drDXS), seem to align properly as the orientation of the subgroups of these residues are similar per couple.

The difference in orientation of Arg478 (ecDXS) and Arg480 (drDXS) show in figures 8 and 9 could be caused by the residues 300-323 loop which is only visible in the drDXS crystal structure and therefore the orientation of this loop in the ecDXS could be different. This difference in orientation could cause a different potency for potential future drugs when it is targeted on DXS enzymes of different origins and this possibility should be considered during future development of such drugs.

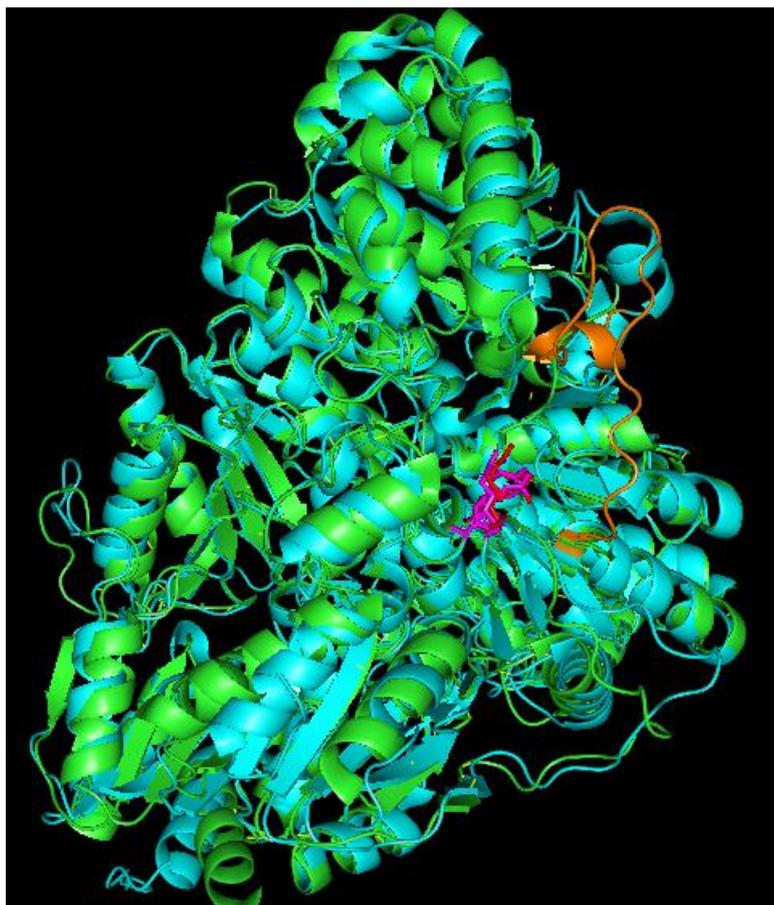


Figure 7: Alignment of the crystal structures of ecDXS (2O1S; green) and drDXS (2O1X; cyan). The Arg420-, Arg478- and Tyr392 residues of ecDXS have been coloured red and the Arg423-, Arg480- and Tyr395 residues of drDXS have been coloured magenta. The loop formed by residues 300-323 of drDXS has been coloured orange. The alignment and visual representation of the crystal structures has been performed using Pymol (developed by Schrödinger). The crystal structures have been obtained from the PDB.

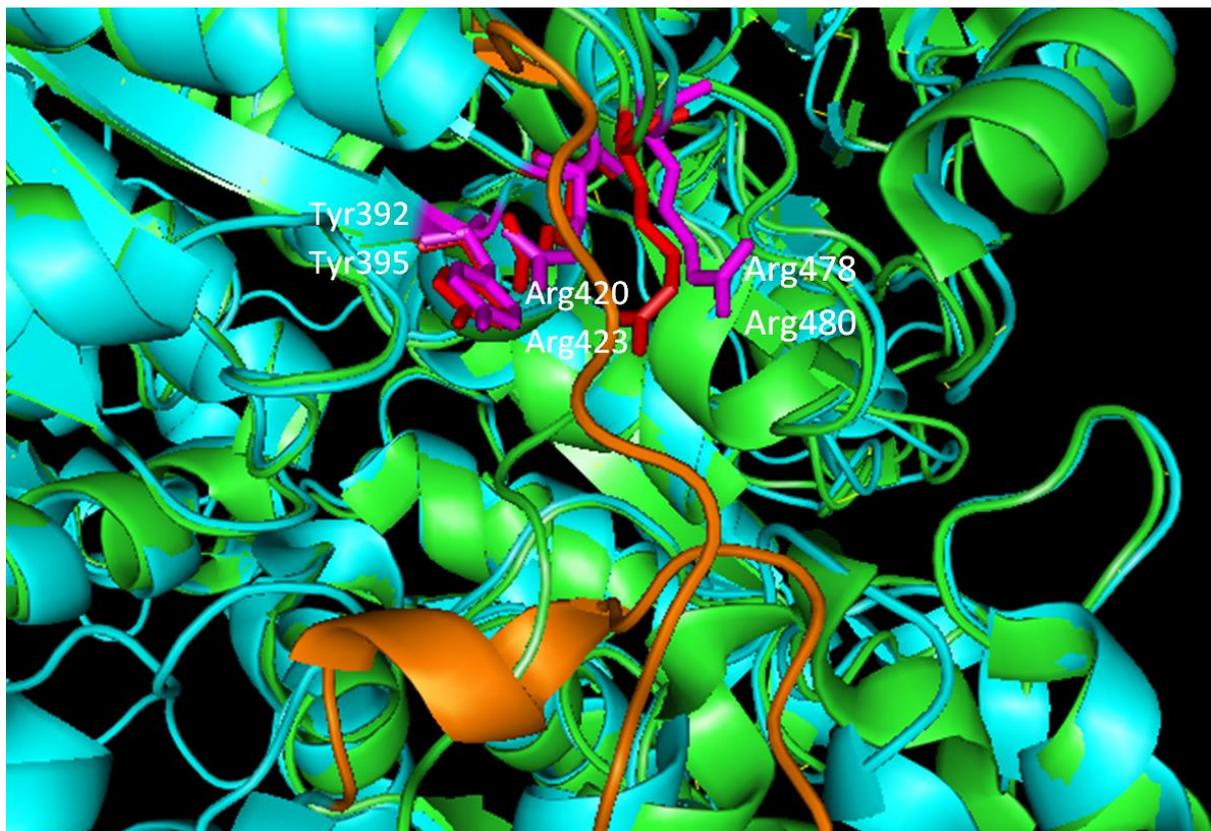


Figure 8: Closer representation of the alignment of ecDXS (201S; green) and drDXS (201X; cyan). The Arg420-, Arg478 and Tyr392 residues of ecDXS (red) were aligned with the Arg423-, Arg480- and Tyr395 residues of drDXS (magenta). The loop formed by residues 300-323 of drDXS has been coloured orange. The alignment and visual representation of the crystal structures has been performed using Pymol (developed by Schrödinger). The crystal structures have been obtained from the PDB.

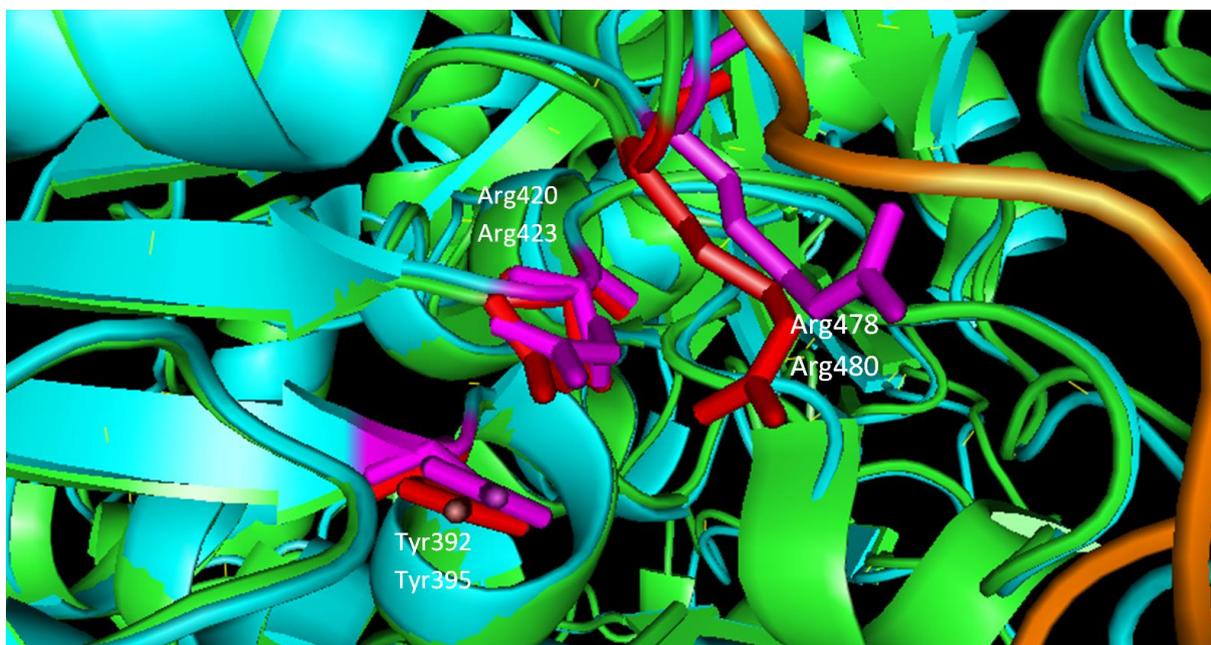


Figure 9: Closer representation of the alignment of ecDXS (201S; green) and drDXS (201X; cyan) from a different angle as in figure yyy. The Arg420-, Arg478 and Tyr392 residues of ecDXS (red) were aligned with the Arg423-, Arg480- and Tyr395 residues of drDXS (magenta). The loop formed by residues 300-323 of drDXS has been coloured orange. The alignment and visual representation of the crystal structures has been performed using Pymol (developed by Schrödinger). The crystal structures have been obtained from the PDB.

1	MS	F--DI	AKYPTLALVDS	TQELRL	L	PKE	SL	PKL	CDELRRY	LLD	SV	SRS	SGH	FASGL	GTVEL	TVAL	HYVYNT	69																																																													
1	MN	E	lpG	TSDT	PLLD	QIHG	PKDLKRL	SREQL	PALTE	ELRGE	IVRVC	SRGGL	H	LASS	LGAVD	II	TAL	HYVLD	71																																																												
1	[358]	FS	[55]	Y	----	F	PLLLK	INN	PSDLK	LK	KQYL	PLLA	HELK	I	FLFF	IV	NIT	GGHFS	SVLS	SLE	IQ	LLLY	IFNQ	478																																																							
1	[11]	II	[53]	Y	s	NR	PT	PLLD	T	INY	PI	HMK	N	SV	KEL	Q	L	SDEL	R	SD	V	F	NV	S	K	T	GGH	L	G	S	S	L	G	V	V	E	L	T	V	A	L	H	I	F	N	T	135																																
1	ML	Q	-----	Q	IRG	PADL	QLH	SQA	QL	REL	AAE	I	REF	L	I	H	K	V	A	A	T	G	H	L	G	P	N	L	G	V	V	E	L	T	L	A	L	H	R	V	F	D	S	60																																			
1	MS	F--DI	AKYPTLALVDS	TQELRL	L	PNE	SL	PKL	CDELRRY	LLD	SV	SRS	SGH	FASGL	GTVEL	TVAL	HYVYNT	69																																																													
1	MS	F--DI	AKYPTLALVDS	TQELRL	L	PKE	SL	PKL	CDELRRY	LLD	SV	SRS	SGH	FASGL	GTVEL	TVAL	HYVYNT	69																																																													
70	PFD	QL	I	W	D	V	G	H	Q	A	Y	P	H	K	I	L	T	G	R	D	K	I	G	T	T	R	Q	K	G	L	H	P	F	P	W	R	G	E	S	E	Y	D	V	L	S	V	G	H	S	S	T	S	I	S	A	G	I	G	I	A	V	A	A	E	K	E	G	-	K	N	R	R	T	V	C	148			
72	PR	DR	I	L	F	D	V	G	H	Q	A	Y	A	H	K	I	L	T	G	R	D	Q	M	A	D	I	K	K	E	G	S	I	S	G	F	T	K	V	S	E	S	E	H	D	A	I	T	V	G	H	A	S	T	S	L	A	N	A	L	G	M	A	L	A	R	D	A	Q	G	-	K	D	F	H	V	A	150		
479	P	Y	D	N	V	I	D	I	G	H	Q	A	Y	V	H	K	I	L	T	G	R	K	L	F	L	S	L	R	N	K	G	I	S	G	F	L	N	I	F	E	S	I	Y	D	K	F	G	A	G	H	S	S	T	S	L	S	A	I	Q	G	Y	E	A	E	W	Q	V	K	N	K	E	Y	---	555					
136	P	Q	D	K	I	L	W	D	V	G	H	Q	A	Y	P	H	K	I	L	T	G	R	G	K	M	P	T	M	R	Q	T	N	G	L	S	G	F	T	K	R	G	E	S	E	H	D	C	F	G	T	G	H	S	S	T	T	I	S	A	G	L	G	M	A	V	G	R	D	L	K	G	-	K	N	N	V	V	A	214
61	P	H	D	P	I	I	F	D	T	G	H	Q	A	Y	V	H	K	M	L	T	G	R	S	Q	D	F	A	T	L	R	K	K	G	L	S	G	Y	P	S	R	A	E	S	E	H	D	W	V	E	S	S	H	A	S	A	L	S	Y	A	D	G	L	A	K	A	F	E	L	T	G	H	R	N	R	H	V	V	A	140
70	PFD	QL	I	W	D	V	G	H	Q	A	Y	P	H	K	I	L	T	G	R	D	K	I	G	T	T	R	Q	K	G	L	H	P	F	P	W	R	G	E	S	E	Y	D	V	L	S	V	G	H	S	S	T	S	I	S	A	G	I	G	I	A	V	A	A	E	K	E	G	-	K	N	R	R	T	V	C	148			
70	PFD	RL	I	W	D	V	G	H	Q	A	Y	P	H	K	I	L	T	G	R	D	K	I	G	T	T	R	Q	K	G	L	H	P	F	P	W	R	G	E	S	E	Y	D	V	L	S	V	G	H	S	S	T	S	I	S	A	G	I	G	V	A	I	A	A	A	K	E	D	-	K	Q	R	R	A	V	C	148			
149	V	I	G	D	G	A	I	T	A	G	M	A	F	E	A	M	N	H	A	G	D	I	R	P	D	M	L	V	I	L	N	D	N	E	M	S	I	S	E	N	V	G	A	L	N	N	H	L	A	Q	L	L	S	G	K	L	Y	S	S	L	R	E	G	G	K	V	F	S	G	V	P	-	P	I	K	223			
151	V	I	G	D	G	S	L	T	G	G	M	A	L	A	L	N	T	I	G	D	M	G	R	K	M	L	I	V	L	N	D	N	E	M	S	I	S	E	N	V	G	A	M	N	K	F	M	R	G	L	Q	V	Q	K	W	F	Q	E	G	E	G	A	G	K	K	A	V	E	A	V	S	K	P	L	A	226			
556	--	C	N	G	I	E	L	S	D	N	A	N	V	T	N	N	E	R	I	F	Q	K	G	---	I	H	N	D	N	N	N	N	I	N	[2]	---	592																																										
215	V	I	G	D	G	A	M	T	A	G	A	Y	E	A	M	N	A	G	L	D	S	I	M	I	V	I	L	N	D	N	K	Q	V	S	L	P	[9]	V	G	A	L	S	S	A	L	S	R	L	Q	S	N	P	A	L	R	E	L	R	E	V	A	K	G	M	T	K	Q	I	C	G	P	M	H	299					
141	V	G	D	G	A	L	T	G	G	M	C	W	E	A	L	N	N	I	A	A	S	R	P	V	I	I	V	V	N	D	N	G	R	S	Y	A	P	T	I	G	G	V	A	D	H	L	A	T	L	R	L	Q	P	A	Y	E	Q	A	L	E	T	G	R	D	L	V	R	A	V	P	-	L	V	[5]	219				
149	V	I	G	D	G	A	I	T	A	G	M	A	F	E	A	M	N	H	A	G	D	I	R	P	D	M	L	V	I	L	N	D	N	E	M	S	I	S	E	N	V	G	A	L	N	N	H	L	A	Q	L	L	S	G	K	L	Y	S	S	L	R	E	G	G	K	V	F	S	G	V	P	-	P	I	K	223			
149	V	I	G	D	G	A	I	T	A	G	M	A	F	E	A	M	N	H	A	G	D	I	K	P	D	L	L	V	I	L	N	D	N	E	M	S	I	S	E	N	V	G	A	L	N	N	H	L	A	Q	L	L	S	G	K	L	Y	S	T	L	R	E	G	G	K	V	F	S	G	V	P	-	P	I	K	223			
224	E	L	L	K	R	T	E	E	H	I	K	G	M	V	P	G	T	---	F	E	E	L	G	F	N	Y	I	G	P	V	D	G	H	D	V	L	G	L	I	T	T	L	K	N	M	R	D	L	---	G	P	Q	F	L	H	I	M	T	K	K	R	G	R	G	Y	E	P	A	E	K	D	294							
227	D	F	M	S	---	R	A	K	N	S	T	R	H	F	D	P	A	S	V	n	p	F	A	A	M	G	V	R	Y	V	G	P	D	G	H	N	V	Q	E	L	V	W	L	L	E	R	L	V	D	L	---	G	P	T	I	L	H	I	V	T	T	K	G	K	L	S	Y	A	E	A	D	299							
593	---	S	R	A	K	N	S	T	R	H	F	D	P	A	S	V	n	p	F	A	A	M	G	V	R	Y	V	G	P	D	G	H	N	T	E	E	L	F	K	V	L	N	N	I	K	E	N	K	l	k	R	A	T	V	L	H	V	R	T	K	S	N	D	F	I	N	S	-	K	S	799								
300	Q	L	A	A	K	V	D	E	Y	A	R	G	M	I	S	G	T	G	S	=	1	F	E	E	L	G	L	Y	I	G	P	V	D	G	H	N	I	D	D	L	V	A	I	L	K	E	V	K	S	T	r	t	G	P	V	L	I	H	V	V	T	E	K	R	G	R	Y	P	A	E	R	A	374						
220	R	F	L	H	S	V	K	A	G	I	K	D	S	L	S	P	Q	L	---	F	T	D	L	G	L	K	Y	V	G	P	V	D	G	H	D	E	R	A	V	E	V	A	L	R	S	A	R	R	F	G	---	A	P	V	I	V	H	V	T	R	K	G	M	G	Y	P	P	A	E	A	D	290							
224	E	L	L	K	R	T	E	E	H	I	K	G	M	V	P	G	T	---	F	E	E	L	G	F	N	Y	I	G	P	V	D	G	H	D	V	L	G	L	I	T	T	L	K	N	M	R	D	L	---	G	P	Q	F	L	H	I	M	T	K	K	R	G	R	G	Y	E	P	A	E	K	D	294							
224	E	L	L	K	R	T	E	E	H	I	K	G	M	V	P	G	T	---	F	E	E	L	G	F	N	Y	I	G	P	V	D	G	H	D	V	L	G	L	V	S	T	L	K	N	M	R	D	L	---	G	P	Q	F	L	H	I	M	T	K	K	R	G	R	G	Y	E	P	A	E	K	D	294							
295	P	I	-	T	F	H	A	V	P	K	F	D	P	S	---	S	G	L	P	K	S	S	G	L	P	S	Y	S	K	I	F	G	D	W	L	C	E	T	A	A	K	D	N	K	I	M	A	I	T	P	A	M	R	E	G	S	G	M	V	E	F	S	R	K	F	P	D	R	Y	F	364								
300	P	I	-	Y	W	H	G	P	A	K	F	D	P	A	---	T	G	E	Y	V	P	S	S	A	---	Y	S	W	S	A	A	F	T	E	A	V	M	E	W	A	K	T	D	P	R	T	F	V	V	T	P	A	M	R	E	G	S	G	L	V	E	F	S	R	V	H	P	H	R	L	Y	367							
761	P	I	=	L	H	S	I	K	[5]	F	D	T	T	=	L	N	G	I	H	K	E	N	K	[35]	S	K	E	T	F	D	I	Y	T	E	N	E	L	F	K	V	L	N	N	I	K	E	N	K	l	k	R	A	T	V	L	H	V	R	T	K	S	N	D	F	I	N	S	-	K	S	873								
375	D	D	-	K	Y	H	G	V	V	K	F	D	P	A	---	T	G	R	Q	F	K	T	T	N	K	T	Q	S	Y	T	T	Y	F	A	E	L	V	A	E	A	E	V	D	K	D	V	V	A	I	H	A	M	G	G	T	G	L	N	L	F	Q	R	R	F	P	T	R	C	F	444									
291	Q	A	=	Q	M	H	S	T	V	P	I	D	P	A	---	T	G	Q	A	T	K	V	A	G	---	P	G	W	T	A	T	F	S	D	A	L	I	G	Y	A	Q	K	R	R	D	I	V	A	I	T	A	A	M	P	G	P	T	G	L	T	A	F	G	Q	R	F	P	D	R	L	F	359							
295	P	I	-	T	F	H	A	V	P	K	F	D	P	S	---	S	G	L	P	K	S	S	G	L	P	S	Y	S	K	I	F	G	D	W	L	C	E	T	A	A	K	D	N	K	I	M	A	I	T	P	A	M	R	E	G	S	G	M	V	E	F	S	R	K	F	P	D	R	Y	F	364								
295	P	I	-	T	F	H	A	V	P	K	F	D	H	T	---	S	G	V	L	P	K	S	S	G	L	P	S	Y	S	K	I	F	G	D	W	L	C	E	T	A	A	K	D	S	K	I	M	A	I	T	P	A	M	R	E	G	S	G	M	V	E	F	S	K	K	F	P	D	Q	Y	F	364							
365	D	V	A	I	A	E	Q	H	A	V	T	F	A	A	G	L	A	I	G	G	-	Y	K	P	V	A	I	V	S	T	F	L	Q	R	A	Y	D	Q	V	L	H	D	V	A	I	Q	K	L	P	V	L	F	A	I	D	R	A	I	V	G	A	D	G	Q	T	H	Q	G	A	F	D								

Sequence analysis for multiple forms of DXS

It has been proposed that DXS enzymes all have a characteristic DRAG sequence (Hahn, Eubanks, Testa, Blagg, Baker, & Poulter, 2001; Bailey, Mahapatra, Brennan, & Crick, 2002), which represents residues 419-422 in ecDXS and thus includes Arg420 (figure 10). Interestingly, this DRAG sequence has not been found in other ThDP dependent enzymes which maintain a classical ordered mechanism or ping-pong mechanism. These observations suggest that the DRAG sequence and Arg420 itself might play an important role in GAP recognition in the rapid equilibrium random sequential mechanism which occurs in DXS enzymes. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014) Another remarkable observation was that the LThDP decarboxylation rate was also accelerated by D-glyceraldehyde. This suggests that the phosphate group of GAP is not necessary for the catalysis of LThDP decarboxylation. (Basta, Patel, Kakalis, Jordan, & Meyers, 2014)

In figure 10 a multiple sequence alignment is shown for DXS enzymes from the microorganisms *E. coli*, *D. radiodurans*, *P. falciparum*, *A. thaliana*, *M. tuberculosis*, *S. flexneri* and *K. pneumoniae*. These organisms have been chosen as they differ in genus. The DRAG sequences and the Tyr392, Arg420, Arg478 and their alignments from the other microorganisms have been highlighted. Interestingly, as these residues are expected to be important for DXS activity (Xiang, Usunow, Lange, Busch, & Tong, 2007; Basta, Patel, Kakalis, Jordan, & Meyers, 2014), the alignment showed a Lys473 for *Mycobacterium tuberculosis* where an Arg473 was expected to be, as is observed for the other microorganisms that the residue is aligned to, however this might not have a large effect on the functionality of this DXS since lysine and arginine are biochemically rather similar. Hahn et al. stated that DXS enzymes all have a characteristic DRAG sequence, but by performing a multiple sequence alignment including *P. falciparum* (figure 10) it is shown that there is no DRAG sequence present in *P. falciparum*. Although the entire DRAG sequence is not represented in *P. falciparum*, the Arg930 was shown to be aligned to Arg420 from *E. coli*, which might further indicate the importance of this residue for the functionality of DXS. The sequence of *P. falciparum* DXS (pfDXS) was shown to have multiple amino acid chains which were unprecedented in the other DXS types and it appeared to lack an amino acid chain which is present in the other DXS types. To attempt to clarify the reason for these differences, the phylogenetic tree was constructed, shown in figure 11.

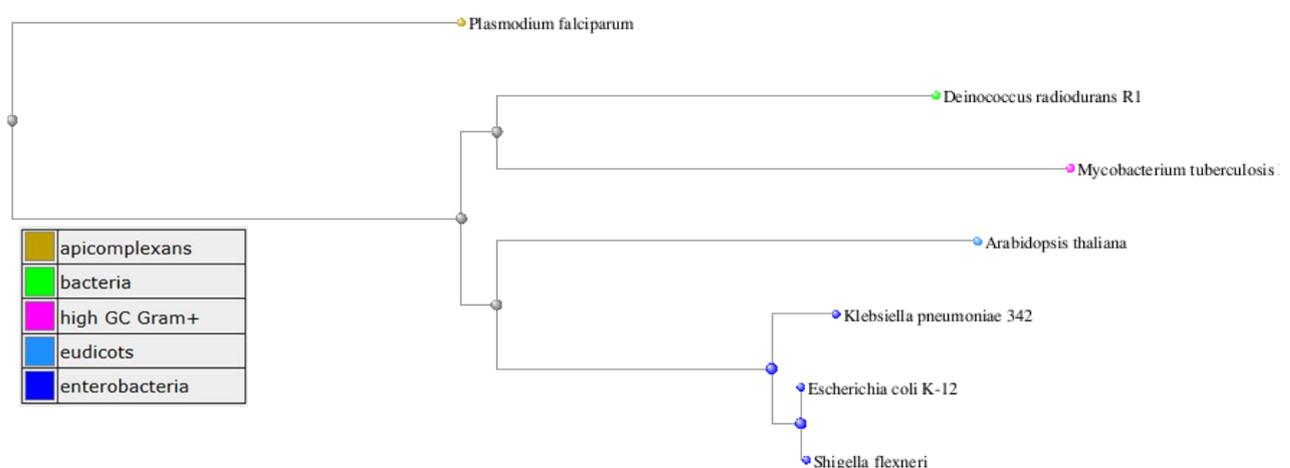


Figure 11: Phylogenetic tree for the DXS types from *E. coli*, *D. radiodurans*, *A. thaliana*, *M. tuberculosis*, *P. falciparum*, *S. flexneri* and *K. pneumoniae*. The sequence used for the construction of the phylogenetic tree were obtained from the NCBI protein sequence database.

The phylogenetic tree shows that *P. falciparum* is a member of the apicomplexans and that it forms one of the two major subgroups, while the other microorganisms form the other major subgroup and are presumably phylogenetically more related to each other than to *P. falciparum*. This could account for the larger diversity in amino acid chains in the DXS sequence between *P. falciparum* and the other microorganisms.

It is noticed that the presumably important residues, aligned to Arg420, Arg478 and Tyr392 of ecDXS, are also present in the pfDXS, even though the sequences from DXS from these two organisms seem very different. The presence of these residues in the pfDXS supports the suggested importance of these residues for the activity of DXS.

Figure 10 shows multiple differences between the sequence of pfDXS and the other DXS types. Residue groups 1-358, 361-405, 600-720, 891-925 and 1017-1060 are not found in the other DXS types. The sequence of *Arabidopsis thaliana* DXS (atDXS), which is a homodimer, is shown to have residues 1-11 and 14-66. Residues 14-66 from atDXS were aligned with residues 361-405 from pfDXS. To inspect whether or not these residue groups are similar, a sequence alignment of atDXS and pfDXS was performed and shown in figure 12, in which the alignment is shown from residues 321-555 and residues 1-142 for pfDXS and atDXS respectively. It is shown that residues 14-66 of atDXS are not aligned with residues 361-405 of pfDXS, which suggests that these groups do not consist of similar amino acid arrangements. Therefore they likely have a different three dimensional structure and might also serve a different purpose within the enzymes. It is also possible that residue groups 1-358 and 361-405 of pfDXS form 1 distinct residue group 1-405, which might serve one or multiple functions. While it has been shown that residues 14-66 of atDXS are not aligned with residues 361-405 of pfDXS, it is likely that residues 1-66 form one distinct residue group.

```

321  NTFINIDEYKTIYGDEIYKEIYELYVERNIPEYYERKYFSEDIKKSVLFDIDKYNDVEFEKAIKEEFINNGVYINNIDNT 400
-----
401  YYKKENILIMKKILHYFPLLKLIINPSDLKCLKKQYLPLLAHELKIFLFFIVNITGGHFSSVLSSEIQLLLLYIFNQPY 480
1    -----MLQQIRGPADLQHLSQLRELAAEIREFLIHKVAATGGHLGPNLGVVELTLALHRVFDSPH 62
481  DNVIIDIGHQAYVHKILTGRKLLFLSLRNKKGISGFLNIFESIYDKFGAGHSSTLSLSAIQGYEAEWQVKNKEY----- 555
63  DPIIFDTGHQAYVHKMLTGRSQDFATLRKKGGLSGYPSRAESEHDWVSSHASAALSADGLAKAFELTGHRNRHVAVV 142

Plasmodium falciparum
Arabidopsis thaliana

```

Figure 12: Sequence alignment of DXS of *P. falciparum* and *A. thaliana*. The sequences were obtained from the NCBI protein sequence database. Total amino acid alignments are coloured red and relatively conserved columns are coloured blue.

Structural modelling *P. falciparum* DXS

P. falciparum has been found to be the most virulent species of the *Plasmodium* genus and most responsible for malaria cases (Rich, Leendertz, Xu, LeBreton, Djoko, Aminake, ..., & Formenty, 2009; Murray, Rosenfeld, Lim, Andrews, Foreman, Haring, ... , & Lopez, 2012) Also *P. falciparum* was found to present resistance to chemotherapeutics and multidrug resistance. (Greenwood, & Mutabingwa, 2002; Petersen, Eastman, & Lanzer, 2011; Park, Lukens, Neafsey, Schaffner, Chang, Valim, ..., & Becker, 2012) It was found that residues

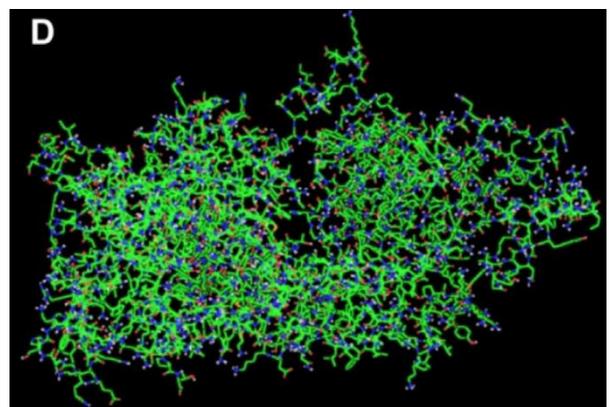


Figure 3: Modelled structure of *P. falciparum* DXS after structural refinement and energy minimisation. The model was developed by Modeller and the image was generated in Pymol (Goswami, 2017)

825 to 971 form the binding pocket for the pyrimidine ring of ThDP and that these residues are well conserved. (Goswami, 2017) Sequence alignments of pfDXS with ecDXS and drDXS show that homology modelling can be carried out for regions 403 to 1205. (Goswami, 2017)

The disordered regions found in pfDXS as compared to ecDXS are important while they can be flexible. This flexibility could cause different possible conformations which could contain important functional sites and therefore could be important for the DXS functionality. (Goswami, 2017) Of twenty five modelled structures generated by Modeller, the best modelled structure has been chosen based on DOPE score and GA341 score. (Goswami, 2017) Subsequently model refinement and energy minimisation have been performed on this structure (figure 13). Another well modelled structure was modelled by RaptorX and it was found that the residues which formed the binding pockets were the same in both models. (Goswami, 2017) The use of these two models showed that many important residues are coinciding for the two models. Using the homology based models developed by RaptorX and Modeller, it was predicted that residues 628 and 657 are involved in Mg(II) binding and residues 855, 877, 879 and 905 are involved in ThDP binding. (Goswami, 2017) Ten compounds were found to be possible candidates for future antimalarial drug design, by screening chemical compounds from ZINC database when using the ThDP binding region of the enzyme. (Goswami, 2017)

Structural modelling *M. tuberculosis* DXS

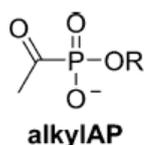
For the generated models of *M. tuberculosis* DXS (mtDXS) ClustalW alignment of the template was used and subsequently the model was improved by energy minimisation. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015) Afterwards DOPE and Verify-3D scores were assessed and Ramachandran plot analyses were carried out. It was found that the model-built mtDXS structure was viable. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015) A high sequence identity was found between drDXS and mtDXS for the ThDP binding pocket and, notably, all the residues in the pocket for binding of the aminopyrimidine ring of ThDP were conserved. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015) Also, the residues in the pocket for binding of the thiazolium ring of ThDP were found to be least conserved, which might indicate a lessened importance of binding of the thiazolium ring for the activity of mtDXS. In drDXS the residues involved in forming hydrophobic interactions with the thiazolium ring are Val80 and Ile187 and together with Ile371 they keep the thiazolium ring situated, but in mtDXS the Val80 and Ile187 have been replaced by Thr69 and Tyr177, which are more hydrophilic and therefore are less successful in forming the hydrophobic interactions with the thiazolium ring. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015) Especially the Tyr177 is situated unfavourable for interaction with the thiazolium ring, while it has an edge-to-face π -stacking interaction with His296 which causes the distance between Tyr177 and the thiazolium ring to be too great to form an interaction. (Masini, Lacy, Monjas, Hawksley, De Voogd, Illarionov, ... , & Hirsch, 2015)

Alkylacetylphosphonates

It has been shown that alkylAPs can inhibit DXS selectively (Smith, Vierling, & Meyers, 2012) and that butylacetylphosphonate (BAP) causes antimicrobial effects through a mechanism which likely involves reversible DXS inhibition. (Smith, Warrington, Vierling, Kuhn, Anderson, Koppisch, & Meyers,

2014) It is also shown that the alkylAPs inhibit DXS more potently than it inhibits other ThDP dependent enzymes such as PDH and TK. (Smith, Vierling, & Meyers, 2012) This selectivity is suggested to be caused by the large active site volume of DXS and its necessity to form a ternary complex. (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017) Straight-chain alkylAPs were found to be reversible, competitive inhibitors and K_i values were found in low micromolar range (table 3).

Table 3: Inhibition of DXS and PDH by alkylAPs (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017)



cpd	R =	K_i^{DXPS} (μM)	K_i^{PDH} (μM)
1 ^a	CH ₃	1.0 ± 0.3	30 ± 10
2 ^a	CH ₂ CH ₃	6.71 ± 0.02	47 ± 2
3	(CH ₂) ₂ CH ₃	7.1 ± 0.8	62 ± 8
4 ^a	(CH ₂) ₃ CH ₃	5.6 ± 0.8	335 ± 8
5	(CH ₂) ₄ CH ₃	1.7 ± 0.3	240 ± 40
6	(CH ₂) ₅ CH ₃	8.9 ± 0.5	117 ± 9
7	(CH ₂) ₇ CH ₃	1.4 ± 0.2	57 ± 5

^aPreviously characterised (Smith, Vierling, & Meyers, 2012)

Also found was a decreased DXS inhibitory activity by the branched isopropylacetylphosphonate (isopropylAP) which has an R = CH(CH₃)₂ side chain. This could indicate that isopropylAP causes unwanted steric interactions with DXS near the binding site of the pyruvate cofactor. Butylacetylphosphonate (butylAP; 4, table 3) was found to be the most potent inhibitor of bacterial growth, which contradicts the expectation that an increase in chain length would result in an increased inhibition of bacterial growth. (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017) As longer chain alkylAPs than butylAP were found to be less potent, it was suggested that this low potency might be caused by a lower permeability or a higher efflux of these compounds. (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017) As only octylacetylphosphonate (octylAP) has been found to be a substrate for the AcrAB-TolC efflux pump, it is suggested that hexylacetylphosphonate (hexylAP) efflux is occurring through a different mechanism or a different pump. Apart from possible differences in efflux between alkylAPs, differences in cellular uptake could also play a significant role in bacterial growth inhibition. Whether the cellular uptake of alkylAPs is mediated by diffusion or active transport is unknown, thus more research is required to understand the cellular uptake of alkylAPs. (Sanders, Vierling, Bartee, DeColli, Harrison, Aklinski, ..., & Freel Meyers, 2017)

Different applications for DXS targeting compounds

DXS has been proposed as a target for novel antibiotics, as well as a target for herbicides. (Banerjee, & Sharkey, 2014) In recent times new antibiotic agents are desired, as antibiotic resistance in pathogens increases. Also antibiotic usage should be minimized and compliance in antibiotic using patients should be optimized. Keeping this in mind raises questions for the applications of future DXS

targeting compounds and whether they should be restricted to antibiotic and antimalarial applications only as opposed to applying them as herbicides. Carefully considering the applications of new compounds is therefore important to minimize the possibility for resistance development against these new compounds.

Increasing artemisinin production by overexpression of *Artemisia annua* DXS

Artemisia annua is an organism which produces the effective antimalarial drug, artemisinin. (Zhang, Liu, Xia, Zeng, Xiang, Zhu, ... , & Liao, 2018) However, the production of artemisinin is insufficient to treat all malaria infected patients and *Plasmodium* species resistant to artemisinin have been reported. (Muangphrom, Seki, Fukushima, & Muranaka, 2016) Results from a study of Zhang et al. suggested that one of the three DXS genes in *A. annua*, *A. annua* DXS2 (AaDXS2) might be the only one of these genes involved in artemisinin biosynthesis. (Zhang, Liu, Xia, Zeng, Xiang, Zhu, ... , & Liao, 2018) As the formation of DXP from DXS is the rate limiting step in the MEP pathway (Kudoh, Hotta, Sekine, Fujii, Uchida, Kubota, ... , & Ihara, 2017; Zhang, Liu, Xia, Zeng, Xiang, Zhu, ... , & Liao, 2018) and the MEP pathway is used in *A. annua* for the biosynthesis of artemisinin, an increased expression of the AaDXS gene in *A. annua* might cause an increase in *Artemisia annua* DXS (AaDXS) which could increase the artemisinin production.

Conclusion

New antibiotic compounds are needed to combat the increasing antibiotic resistance developed in pathogenic bacteria and therefore new targeting mechanisms are researched. DXS remains an interesting target for new antibiotics and antimalarials, as DXS is present in many organisms. Because the DXS gene is expressed in a wide variety of organisms, it is expected that there are many interspecies differences between the enzyme variants. This is also shown in figure 10, where the sequence of pfDXS deviated from the DXS sequences from the other organisms it was aligned to. Therefore, the diversity of DXS variants should be taken into account when developing new DXS targeting compounds. Genetic and structural studies could provide vital information about this diversity of DXS variants and about possible residues which could be targeted. In particular identification of residues involved in the working mechanism is favourable, as these residues are expected to be more similar between DXS variants and thus these residues could better act as a target for a universal DXS targeting compound.

References

- Bailey, A. M., Mahapatra, S., Brennan, P. J., & Crick, D. C. (2002). Identification, cloning, purification, and enzymatic characterization of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate synthase. *Glycobiology*, *12*(12), 813-820.
- Banerjee, A., Preiser, A. L., & Sharkey, T. D. (2016). Engineering of recombinant poplar deoxy-D-xylulose-5-phosphate synthase (PtDXS) by site-directed mutagenesis improves its activity. *PloS one*, *11*(8), e0161534.
- Banerjee, A., & Sharkey, T. D. (2014). Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Natural product reports*, *31*(8), 1043-1055.
- Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H., & Sharkey, T. D. (2013). Feedback inhibition of deoxy-D-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *Journal of Biological Chemistry*, *288*(23), 16926-16936.
- Bartee, D., & Freil Meyers, C. L. (2018). Targeting the unique mechanism of bacterial 1-deoxy-D-xylulose-5-phosphate synthase. *Biochemistry*, *57*(29), 4349-4356.
- Basta, L. A. B., Patel, H., Kakalis, L., Jordan, F., & Meyers, C. L. F. (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. *The FEBS journal*, *281*(12), 2820-2837.
- Battistini, M. R., Shoji, C., Handa, S., Breydo, L., & Merkler, D. J. (2016). Mechanistic binding insights for 1-deoxy-D-Xylulose-5-Phosphate synthase, the enzyme catalyzing the first reaction of isoprenoid biosynthesis in the malaria-causing protists, *Plasmodium falciparum* and *Plasmodium vivax*. *Protein expression and purification*, *120*, 16-27.
- Brammer, L. A., Smith, J. M., Wade, H., & Meyers, C. L. F. (2011). 1-Deoxy-D-xylulose 5-phosphate synthase catalyzes a novel random sequential mechanism. *Journal of Biological Chemistry*, *286*(42), 36522-36531.
- Crowell, D. N., Packard, C. E., Pierson, C. A., Giner, J. L., Downes, B. P., & Chary, S. N. (2003). Identification of an allele of CLA1 associated with variegation in *Arabidopsis thaliana*. *Physiologia plantarum*, *118*(1), 29-37.
- DeColli, A. A., Nemeria, N. S., Majumdar, A., Gerfen, G. J., Jordan, F., & Meyers, C. L. F. (2018). Oxidative decarboxylation of pyruvate by 1-deoxy-d-xylulose 5-phosphate synthase, a central metabolic enzyme in bacteria. *Journal of Biological Chemistry*, *293*(28), 10857-10869.
- Estévez, J. M., Cantero, A., Romero, C., Kawaide, H., Jiménez, L. F., Kuzuyama, T., ... & León, P. (2000). Analysis of the expression of CLA1, a gene that encodes the 1-deoxyxylulose 5-phosphate synthase of the 2-C-methyl-D-erythritol-4-phosphate pathway in *Arabidopsis*. *Plant physiology*, *124*(1), 95-104.
- Frank, A., & Groll, M. (2016). The methylerythritol phosphate pathway to isoprenoids. *Chemical reviews*, *117*(8), 5675-5703.
- Frank, R. A. W., Leeper, F. J., & Luisi, B. F. (2007). Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cellular and Molecular Life Sciences*, *64*(7-8), 892.

- Goswami, A. M. (2017). Computational analysis, structural modeling and ligand binding site prediction of Plasmodium falciparum 1-deoxy-d-xylulose-5-phosphate synthase. *Computational biology and chemistry*, 66, 1-10.
- Greenwood, B., & Mutabingwa, T. (2002). Malaria in 2002. *Nature*, 415(6872), 670.
- Hahn, F. M., Eubanks, L. M., Testa, C. A., Blagg, B. S., Baker, J. A., & Poulter, C. D. (2001). 1-Deoxy-d-Xylulose 5-Phosphate Synthase, the Gene Product of Open Reading Frame (ORF) 2816 and ORF 2895 in Rhodobacter capsulatus. *Journal of bacteriology*, 183(1), 1-11.
- Hailes, H. C., Rother, D., Müller, M., Westphal, R., Ward, J. M., Pleiss, J., ... & Pohl, M. (2013). Engineering stereoselectivity of ThDP-dependent enzymes. *The FEBS journal*, 280(24), 6374-6394.
- Handa, S., Dempsey, D. R., Ramamoorthy, D., Cook, N., Guida, W. C., Spradling, T. J., ... & Merkle, D. J. (2018). Mechanistic Studies of 1-Deoxy-D-Xylulose-5-Phosphate Synthase from Deinococcus radiodurans. *Biochemistry & molecular biology journal*, 4(1).
- Hayashi, D., Kato, N., Kuzuyama, T., Sato, Y., & Ohkanda, J. (2013). Antimicrobial N-(2-chlorobenzyl)-substituted hydroxamate is an inhibitor of 1-deoxy-D-xylulose 5-phosphate synthase. *Chemical Communications*, 49(49), 5535-5537.
- Kudoh, K., Hotta, S., Sekine, M., Fujii, R., Uchida, A., Kubota, G., ... & Ihara, M. (2017). Overexpression of endogenous 1-deoxy-d-xylulose 5-phosphate synthase (DXS) in cyanobacterium Synechocystis sp. PCC6803 accelerates protein aggregation. *Journal of bioscience and bioengineering*, 123(5), 590-596.
- Masini, T., Lacy, B., Monjas, L., Hawksley, D., De Voogd, A. R., Illarionov, B., ... & Hirsch, A. K. H. (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. *Organic & biomolecular chemistry*, 13(46), 11263-11277.
- Matsue, Y., Mizuno, H., Tomita, T., Asami, T., Nishiyama, M., & Kuzuyama, T. (2010). The herbicide ketoclozazole 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. *The Journal of antibiotics*, 63(10), 583.
- Muangphrom, P., Seki, H., Fukushima, E. O., & Muranaka, T. (2016). Artemisinin-based antimalarial research: application of biotechnology to the production of artemisinin, its mode of action, and the mechanism of resistance of Plasmodium parasites. *Journal of natural medicines*, 70(3), 318-334.
- Mueller, C., Schwender, J., Zeidler, J., & Lichtenthaler, H. K. (2000). Properties and inhibition of the first two enzymes of the non-mevalonate pathway of isoprenoid biosynthesis.
- Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., ... & Lopez, A. D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *The Lancet*, 379(9814), 413-431.
- Park, D. J., Lukens, A. K., Neafsey, D. E., Schaffner, S. F., Chang, H. H., Valim, C., ... & Becker, J. S. (2012). Sequence-based association and selection scans identify drug resistance loci in the Plasmodium falciparum malaria parasite. *Proceedings of the National Academy of Sciences*, 109(32), 13052-13057.
- Petersen, I., Eastman, R., & Lanzer, M. (2011). Drug-resistant malaria: Molecular mechanisms and implications for public health. *FEBS letters*, 585(11), 1551-1562.

- Pulido, P., Llamas, E., Llorente, B., Ventura, S., Wright, L. P., & Rodríguez-Concepción, M. (2016). Specific Hsp100 chaperones determine the fate of the first enzyme of the plastidial isoprenoid pathway for either refolding or degradation by the stromal Clp protease in *Arabidopsis*. *PLoS genetics*, *12*(1), e1005824.
- Pulido, P., Toledo-Ortiz, G., Phillips, M. A., Wright, L. P., & Rodríguez-Concepción, M. (2013). *Arabidopsis* J-protein J20 delivers the first enzyme of the plastidial isoprenoid pathway to protein quality control. *The Plant Cell*, *25*(10), 4183-4194.
- Querol, J., Rodríguez-Concepción, M., Boronat, A., & 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. *Biochemical and biophysical research communications*, *289*(1), 155-160.
- Querol-Audí, J., Boronat, A., Centelles, J. J., & Imperial, S. (2014). Catalytically important residues in *E. coli* 1-deoxy-D-xylulose 5-phosphate synthase. *Journal of Biosciences and Medicines*, *2*(04), 30.
- Ramamoorthy, D., Handa, S., Merkle, D. J., & Guida, W. C. (2014). Plasmodium Vivax 1-Deoxy-D-Xylulose-5-Phosphate Synthase: Homology Modeling, Domain Swapping, and Virtual Screening. *Journal of Data Mining in Genomics & Proteomics*, (S1), 1.
- Rich, S. M., Leendertz, F. H., Xu, G., LeBreton, M., Djoko, C. F., Aminake, M. N., ... & Formenty, P. (2009). The origin of malignant malaria. *Proceedings of the National Academy of Sciences*, *106*(35), 14902-14907.
- Sanders, S., Vierling, R. J., Barte, D., DeColli, A. A., Harrison, M. J., Aklinski, J. L., ... & Freel Meyers, C. L. (2017). Challenges and Hallmarks of Establishing Alkylacetylphosphonates as Probes of Bacterial 1-Deoxy-d-xylulose 5-Phosphate Synthase. *ACS infectious diseases*, *3*(7), 467-478.
- Smith, J. M., Vierling, R. J., & Meyers, C. F. (2012). Selective inhibition of *E. coli* 1-deoxy-D-xylulose-5-phosphate synthase by acetylphosphonates. *MedChemComm*, *3*(1), 65-67.
- Smith, J. M., Warrington, N. V., Vierling, R. J., Kuhn, M. L., Anderson, W. F., Koppisch, A. T., & Meyers, C. L. F. (2014). Targeting DXP synthase in human pathogens: enzyme inhibition and antimicrobial activity of butylacetylphosphonate. *The Journal of antibiotics*, *67*(1), 77.
- Xiang, S., Usunow, G., Lange, G., Busch, M., & Tong, L. (2007). Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *Journal of Biological Chemistry*, *282*(4), 2676-2682.
- Zhang, F., Liu, W., Xia, J., Zeng, J., Xiang, L., Zhu, S., ... & Liao, Z. (2018). Molecular characterization of the 1-deoxy-D-xylulose 5-phosphate synthase gene family in *Artemisia annua*. *Frontiers in plant science*, *9*.
- Zhao, L., Chang, W. C., Xiao, Y., Liu, H. W., & Liu, P. (2013). Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annual review of biochemistry*, *82*, 497-530.