Thermodynamic Effects in Enzyme Regulation, Stereochemistry and Process Control

Marsden, S.R.

DOI
10.4233/uuid:3353f734-2d23-4dbd-b80d-ffac899c69e8

Publication date
2021

Document Version
Final published version

Citation (APA)
Marsden, S. R. (2021). *Thermodynamic Effects in Enzyme Regulation, Stereochemistry and Process Control*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:3353f734-2d23-4dbd-b80d-ffac899c69e8

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Thermodynamic Effects in Enzyme Regulation, Stereochemistry and Process Control

Dissertation

for the purpose of obtaining the degree of doctor

at Delft University of Technology

by the authority of the Rector Magnificus, Prof.dr.ir. T.H.H.J. van der Hagen

chair of the Board for Doctorates

to be defended publicly on

Friday 23 April 2021 at 10:00 o'clock

by

Stefan Robert MARSDEN

Master of Science in Life Science and Technology, Delft University of Technology, the Netherlands, born in Baden, Switzerland

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof.dr. U. Hanefeld	Delft University of Technology, promotor
Dr. D.G.G. McMillan	Delft University of Technology, copromotor

Independent members:

Prof.dr. W.R. Hagen	Delft University of Technology, The Netherlands
Prof.dr. H.C. Hailes	University College London, United Kingdom
Prof.dr. U.T. Bornscheuer	University of Greifswald, Germany
Prof.dr. M. Pohl	Forschungszentrum Jülich, Germany
Prof.dr. D.B. Janssen	University of Groningen, The Netherlands
Prof.dr. F. Hollmann	Delft University of Technology, reserve member

Keywords: thermodynamics, kinetic control, thiamine diphosphate, aldolase

Printed by: Gildeprint, Enschede

Cover by: Stefan Robert Marsden

ISBN: XXX

An electronic copy of this dissertation is available at http://repository.tudelft.nl/

to my grandfather

Dipl.-Kfm. Dr. Gisbert Steinhardt

Mit Deiner Freude am Lernen warst du mir von klein auf an Vorbild und Motivation auf meinem schulischen Weg.

Table of contents

Summary		**6**
Samenvatting		**10**
1.	Preface	**15**
2.	Separating Thermodynamics from Kinetics – A New Understanding of the Transketolase Reaction	**19**
3.	Assessing the ThDP-dependent *Ec*PDH E1 subunit for carboligation reactions with aliphatic ketoacids	**35**
4.	CH-π interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase	**61**
5.	Substrate Induced Movement of the Metal Cofactor between Active and Resting State Configurations in a Metalloenzyme	**89**
6.	Investigation of an Aldolase/Transaminase Cascade Reaction for the Synthesis of C-3 Modified Amino Aids	**113**
7.	Thermodynamically and Kinetically Controlled Reactions In Biocatalysis – From Concepts to Perspectives	**149**
8.	Thermodynamics Determine the Diastereochemical Outcome of Catalytic Reactions	**179**
9.	Conclusions and Outlook	**199**
	Acknowledgements	**202**
	Curriculum vitae	**205**
	Publications and patents	**206**

Summary

Thiamine diphosphate dependent enzymes are excellent catalysts for the asymmetric synthesis of the α-hydroxyketone (acyloin) structural motif, which is found in many pharmaceuticals and fine chemicals. In **chapter 2**, variants of transketolase from *Saccharomyces cerevisiae* were screened for the conversion of aliphatic aldehydes with hydroxypyruvate as donor substrate. The formation of a new hydrogen bond network was observed in the most successful variant D477E, which allowed for the accommodation of hydrophobic aldehydes within the enzyme's polar active site. Decarboxylation of hydroxypyruvate was shown to render the carboligation reaction kinetically controlled, correcting the preceding notion of an irreversible conversion of substrates in literature.

Thiamine diphosphate dependent enzymes are notorious for their strict selectivity towards their natural donor substrate, which limits their synthetic product scope. In **chapter 3**, the pyruvate dehydrogenase E1 subunit from *Escherichia coli* (*Ec*PDH E1) was explored as an alternative to the engineering of transketolase for the conversion of aliphatic ketoacids. Here, wild-type *Ec*PDH E1 displayed a 180- to 250-fold higher specific activity towards pyruvate and 2-oxobutyrate, next to a broad acceptor substrate scope, which warrants its application for synthesis. The large hydration constants of electropositive aldehydes were shown to limit the alternate, thermodynamically controlled coupling of aldehydes unviable in aqueous solution. Additionally, a loss of stereocontrol was observed for the thermodynamically controlled approach. The decarboxylation of ketoacids under kinetically controlled conditions is therefore a prerequisite for the continued expansion of the donor substrate scope in thiamine diphosphate dependent enzymes.

However, the chemical synthesis of functionalized ketoacids is laborious, and only few, primarily aliphatic compounds are commercially available. The discovery of a hydroxy ketoacid aldolase from *Sphingomonas wittichii* RW1 (*Sw*HKA) promised to address this issue, by catalysing the modular synthesis of structurally diverse hydroxy ketoacids in a single step. In **chapter 4**, *Sw*HKA was therefore biochemically characterised and its crystal structure was solved. The natural mutation of an otherwise conserved leucine residue L210F was identified to newly establish CH-π interactions with hydroxypyruvate as donor substrate, which promotes its activation to the enol intermediate. A variant of enhanced activity was obtained by further increasing the electron density of the aromatic system in variant F210Y. Notably, *Sw*HKA displayed a micromolar affinity towards inorganic phosphate, which increased its activity up to 120-fold. In this chapter, CH-π interactions were demonstrated to be an important mode of substrate recognition and activation beyond the well-known examples of carbohydrate-active enzymes. Additionally, the function of inorganic phosphate as an unusual cofactor was demonstrated.

In **chapter 5**, an in-depth analysis of crystal structures of *Sw*HKA at up to 1.2 Å resolution revealed a dynamic movement of the metal cofactor between two metal binding sites of distinct properties. In its resting state configuration, the metal cofactor is coordinated at the dimer interface of the *holo*-hexamer, but does not allow for substrate binding. Conversely, a second coordination sphere corresponds to the catalytically active state at 2.4 Å distance. Here, bidentate substrate coordination affords the overall lowest energy complex and constitutes the driving force for the transition of the metal cofactor from its resting configuration into its catalytically active state. This activation mechanism is conceptually distinct from allosteric regulation, as it does not require a concomitant change in protein structure. This observation could serve as a design principle for artificial metalloenzymes with potential applications in systems biocatalysis.

In **chapter 6**, a bienzymatic aldolase-transaminase based cascade reaction was investigated for the synthesis of non-natural, polyhydroxylated amino acids. This bienzymatic approach offers the advantage of complete stereocontrol at three chiral centres (C2-C4). In this study, 102 transaminases from the UCL culture collection were screened for the conversion of D-ketogluconate and D-glucosaminic acid as model compounds for aldol reactions with *Sw*HKA. While low levels of conversion were observed for the thermodynamically preferred conversion of D-glucosaminic acid to D-ketogluconic acid, none of the screened enzymes displayed viable kinetic properties towards these non-natural substrates. The directed evolution of a suitable transaminase scaffold will therefore be required for the synthesis of non-natural, polyhydroxylated amino acids in aldolase-transaminase based cascade reactions with *Sw*HKA.

While water is generally considered a green solvent, it can also negatively affect the product yields of enzymatic reactions. In **chapter 7**, the elimination of a good leaving group to render biocatalytic conversions kinetically controlled was reviewed as a strategy to improve product yields. While the different aspects of thermodynamically- and kinetically controlled syntheses of β-lactam antibiotics with amidohydrolases are well-investigated, this is far less the case for other enzyme classes. Here, the general features of both approaches were summarised at the example of amidohydrolases and conceptually portrayed for four other enzyme classes: acyl transferases, ThDP-dependent enzymes, glycosyl transferases and kinases. Finally, parallels were drawn between qualitative energy diagrams of Diels-Alder reactions and enzyme catalysed conversions. Their similarity led to the notion, that enzymes (and catalysts in general) could catalyse the epimerization of a chiral centre within diastereomers against their intrinsic diastereoselectivity.

Chapter 8 provides extensive experimental and theoretical evidence for the simultaneous occurrence of thermodynamic epimerization during the catalytic synthesis of diastereomers. A multitude of parameters that govern thermodynamic epimerization were investigated at the example of aldol reactions with fluoropyruvate, using two stereocomplementary aldolases *Sw*HKA and *Bp*HKA. Next to low temperatures, a high maximal conversion was identified to promote kinetically controlled conditions. For bimolecular reactions, this parameter can conveniently be modulated *via* the use of one substrate in excess. Quantum mechanical calculations were demonstrated to accurately predict the identity of the thermodynamic product and its associated diastereomeric excess under equilibrium conditions. The common practice of deducing a catalyst's stereoselectivity based on the absolute configuration of the isolated product therefore constitutes a potential pitfall, and warrants the introduction of a new standard for catalyst characterisation.

Samenvatting

Thiaminedifosfaat afhankelijke enzymen zijn uitstekende katalysatoren voor de asymmetrische synthese van het α-hydroxyketon (acyloïne) structurele motief, dat in veel farmaceutische producten en fijnchemicaliën wordt aangetroffen. In **hoofdstuk 2** werden varianten van transketolase uit *Saccharomyces cerevisiae* gescreend op de omzetting van alifatische aldehyden met hydroxypyruvaat als donorsubstraat. De vorming van een nieuw waterstofbindingsnetwerk werd waargenomen in de meest succesvolle variant D477E, die de accommodatie van hydrofobe aldehyden binnen het polaire actieve centrum van het enzym mogelijk maakte. Aangetoond werd dat decarboxylering van hydroxypyruvaat de carboligatiereactie kinetisch gecontroleerd maakt, waardoor het voorgaande idee van een onomkeerbare omzetting voor het aldehydesubstraat in de literatuur werd gecorrigeerd.

Toch zijn thiaminedifosfaat afhankelijke enzymen ook berucht vanwege hun strikte selectiviteit ten opzichte van hun natuurlijke donorsubstraat, wat hun synthetische productomvang beperkt. In **hoofdstuk 3** werd de pyruvaatdehydrogenase E1-subeenheid van *Escherichia coli* (*Ec*PDH E1) onderzocht als alternatief voor de engineering van transketolase voor de omzetting van alifatische ketozuren. Hier vertoonde wild-type *Ec*PDH E1 een 180 tot 250 maal hogere specifieke activiteit ten opzichte van pyruvaat en 2-oxobutyraat, naast een brede acceptorsubstraat, wat de toepassing ervan voor synthese rechtvaardigt. De grote hydratatieconstanten van elektropositieve aldehyden bleken de alternatieve, thermodynamisch gecontroleerde koppeling van aldehyden in waterige oplossing te beperken. Bovendien werd een verlies van stereocontrole waargenomen voor de thermodynamisch gecontroleerde reacties. De decarboxylering van ketozuren onder kinetisch gecontroleerde omstandigheden is daarom een voorwaarde voor de voortdurende uitbreiding van het donorsubstraat in thiaminedifosfaat afhankelijke enzymen.

De chemische synthese van gefunctionaliseerde ketozuren is echter bewerkelijk en er zijn slechts enkele, voornamelijk alifatische verbindingen in de handel verkrijgbaar. De recente ontdekking van een hydroxy ketozuur aldolase van *Sphingomonas wittichii* RW1 (*Sw*HKA) beloofde dit probleem aan te pakken door de modulaire synthese van structureel diverse hydroxy ketozuren in een enkele stap. In **hoofdstuk 4** werd *Sw*HKA daarom biochemisch gekarakteriseerd en werd de kristalstructuur opgelost. De natuurlijke mutatie van een anderszins geconserveerde leucine L210F werd geïdentificeerd om nieuw CH-π-interacties tot stand te brengen met hydroxypyruvaat als donorsubstraat, wat de activering ervan naar het enol-tussenproduct bevordert. Een variant van verhoogde activiteit werd verkregen door de elektronendichtheid van het aromatische systeem in variant F210Y verder te verhogen. Met name vertoonde SwHKA een micromolaire affiniteit voor anorganisch fosfaat, waardoor de activiteit tot 120 keer toenam. In dit hoofdstuk werd aangetoond dat CH-π-interacties een belangrijke manier zijn voor substraatherkenning en activering buiten de bekende voorbeelden van koolhydraat-actieve enzymen. Bovendien werd het gebruik van anorganisch fosfaat als ongebruikelijke cofactor aangetoond.

In **hoofdstuk 5** onthulde een diepgaande analyse van kristalstructuren van *Sw*HKA met een resolutie tot 1,2 Å een dynamische beweging van de metaalcofactor tussen twee metaalbindingsplaatsen met verschillende eigenschappen. In zijn rusttoestand wordt de metalen cofactor gecoördineerd op het dimeer grensvlak van de holo-hexameer, maar is substraatbinding niet mogelijk. Omgekeerd komt een tweede coördinatiesfeer met een lagere bezetting op 2,4 Å afstand overeen met de katalytisch actieve toestand. Hier levert bidentate substraatcoördinatie het algehele laagste energiecomplex op en vormt de drijvende kracht voor de overgang van de metaalcofactor van zijn rustconfiguratie naar zijn katalytisch actieve toestand. Dit mechanisme is conceptueel verschillend van allosterische regulatie, aangezien er geen gelijktijdige verandering van de

eiwitstructuur nodig is. Dit zou kunnen dienen als een ontwerpprincipe voor kunstmatige metallo-enzymen met mogelijke toepassingen in systeembiokatalyse.

In **hoofdstuk 6** werd een bienzymatische aldolase-transaminase cascadereactie onderzocht voor de synthese van niet-natuurlijke, polyhydroxyleerde aminozuren. Deze bienzymatische benadering biedt het voordeel van volledige stereocontrole in drie chirale centra (C2-C4). In deze studie werden 102 transaminasen uit de UCL-cultuurcollectie gescreend op de omzetting van D-ketogluconaat en D-glucosaminezuur als modelverbindingen voor aldolreacties met *Sw*HKA. Hoewel lage omzettingsniveaus werden waargenomen voor de thermodynamisch geprefereerde omzetting van D-glucosaminezuur in D-ketogluconzuur, vertoonde geen van de gescreende enzymen voldoende kinetische eigenschappen ten opzichte van deze niet-natuurlijke substraten. De gerichte evolutie van een geschikte transaminase-scaffold zal daarom vereist zijn voor de synthese van niet-natuurlijke, polyhydroxyleerde aminozuren in aldolase-transaminase cascadereacties met *Sw*HKA.

Hoewel water over het algemeen als een groen oplosmiddel wordt beschouwd, kan het ook de productopbrengst van enzymatische reacties negatief beïnvloeden. In **hoofdstuk 7** werd de eliminatie van een goede afgangsgroep om biokatalytische conversies kinetisch gecontroleerd te maken besproken als een strategie, om de productopbrengst te verbeteren. Hoewel de verschillende aspecten van thermodynamisch en kinetisch gecontroleerde syntheses van β-lactam antibiotica met amidohydrolasen goed onderzocht zijn, is dit veel minder het geval voor andere enzymklassen. Hier werden de algemene kenmerken van beide benaderingen samengevat in het voorbeeld van amidohydrolases en conceptueel overgedragen op vier andere enzymklassen: acyltransferases, ThDP-afhankelijke enzymen, glycosyltransferases en kinasen. Ten slotte werden parallellen getrokken tussen kwalitatieve energiediagrammen van Diels-Alder-reacties en enzymgekatalyseerde conversies. Hun gelijkenis leidde tot de

hypothese dat enzymen (en ook katalysatoren in het algemeen) de inversie van een chiraal centrum zouden kunnen katalyseren tegen hun intrinsieke stereoselectiviteit.

Hoofdstuk 8 levert uitgebreid experimenteel en theoretisch bewijs voor onze hypothese, dat thermodynamische epimerisatie gelijktijdig plaatsvindt tijdens de katalytische synthese van diastereomeren. Een veelvoud aan parameters die de thermodynamische epimerisatie bepalen, werden onderzocht aan de hand van aldolreacties met fluorpyruvaat, met behulp van de twee stereocomplementaire aldolases *Sw*HKA en *Bp*HKA. Naast lage temperaturen werd een hoge maximale conversie geïdentificeerd om kinetisch gecontroleerde omstandigheden te bevorderen. Voor bimoleculaire reacties kan deze parameter gemakkelijk worden gemoduleerd door het gebruik van één substraat in overmaat. Kwantummechanische berekeningen werden uitgevoerd om de identiteit van het thermodynamische product en de bijbehorende diastereomere overmaat onder evenwichtsomstandigheden nauwkeurig te kunnen voorspellen. De gangbare praktijk om de stereoselectiviteit van een katalysator af te leiden uit de absolute configuratie van het geïsoleerde product vormt daarom een mogelijke valkuil voor katalyse, die de introductie van een nieuwe standaard voor de karakterisering van katalysatoren rechtvaardigt.

1

Preface

From a purely scientific perspective, the questions of whether the exact number of leaves in any given tree has changed overnight, and how to cure a potentially fatal disease are equally worthy to ask. It is only within the ideological context of a society, that we can ascertain to which question we give preference and then allocate our limited resources accordingly.

The United Nations were formed as a global organisation to promote peace and prosperity amongst all nations. At the beginning of the 21st century, leaders from 189 countries formulated eight Millennium Development Goals in response to the most urgent societal issues:[1]

Figure 1: Official logos of the eight Millennium Development Goals by the UN.[1]

Many of these issues will require tremendous societal efforts in order to overcome their underlying cultural and political origins. Amongst the multiplicity of required measures, innovation will undoubtedly constitute a cornerstone of the collective response to these challenges, by providing technological solutions.

Here, biotechnology presents itself as a particularly promising discipline.

For instance, genetically engineered crops allow for more reliable, more nutritious and increased harvests, while simultaneously reducing the environmental impact that is caused by the use of fertilisers and pesticides.[2] Further, a fundamental understanding of cellular processes and their genetic blueprints could pave the way towards personalised medicine. This could not only improve the efficacy of established treatments and promote the development of new drugs, but also remedy the pre-existing gender and ethnicity bias in medical research.[3]

Enzymes are nature's favourite tools to create and sustain life. In consequence, the subfield of *Enzyme Engineering* has specialised in the development of functional proteins as essential instruments for diverse biotechnological applications.[4a-e] The unique features of enzymes are derived from a truly modular design: simple variation of four letters (ATGC) in the genetic code allows for an unparalleled control over the three dimensional space on the molecular level.[5a,b] A fundamental understanding of structure-function relationships in proteins and their mechanistic mode of operation therefore holds the promise of tailor-made enzymes for almost any conceivable, molecular problem in biotechnology.[6a] Here, the recent breakthrough in protein structure prediction by AlphaFold can be considered a quantum leap towards this goal.[7]

Yet, much still remains unknown about protein function. In this thesis, different enzymatic modes of operation were investigated for aldolases and thiamine diphosphate dependent enzymes, with a particular focus on the role of thermodynamics. While the herein obtained results only constitute a tiny piece within an enormously sized jigsaw puzzle, we are privileged to watch it slowly take shape.

References

[1] https://www.un.org/millenniumgoals/ (05.10.2020)

[2] https://science.sciencemag.org/content/290/5490/253.full (07.10.2020)

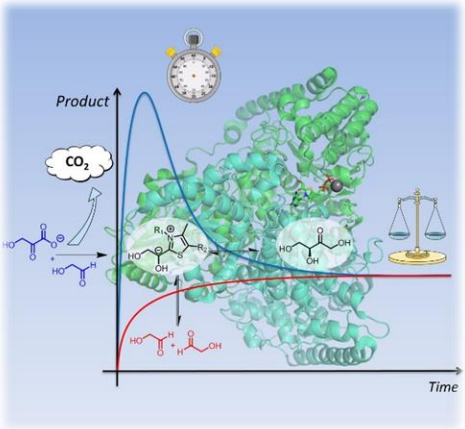
[3] M. W. Nielsen, J. P. Andersen, L. Schiebinger, J. W. Schneider. One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis. *Nat. Hum. Behav.* **2017**, *1*, 791-796.

[4] a) P. Schindele*,* H. Puchta. Engineering CRISP/*Lb*Cas12a for highly efficient, temperature-tolerant plant gene editing. *Plant Biotechnol. J.,* **2019**, *18,* 1118-1120. b) J. Aschenbrenner*,* S. Werner, V. Marchand, M. Adam, Y. Motorin, M. Helm, A. Marx. Engineering of DNA polymerase for direct m^6A sequencing. *Angew. Chem. Int. Ed.*, **2018**, *57*, 417-421. c) M. Akbarian, R. Yousefi, A. Moosavi-Movahedi, A. Ahmad, V. N. Uversky. Modulating insulin fibrillation using engineered B-chains with mutated C-termini. *Biophys. J.,* **2019**, *117*, 1626-1641. d) S. Jiao, F. Li, H. Yu, Z. Shen. Adances in acrylamide bioproduction catalyzed with *Rhodococcus* cells harboring nitrile hydratase. *Appl. Microbio. Biot.* **2019**, *104,* 1001-1012. e), S. H. Kung, S. Lund, A. Murarka, D. McPhee, C. J. Paddon. Approaches and recent developments for the commercial production of semi-synthetic artemisinin. *Front. Plant Sci.* **2018**, *9*, 87.

[5] a) T. Bitard-Feidel, *bioRxiv*, **2020**. https://doi.org/10.1101/2020.11.09.375311. b) P. C. F. Buchholz, C. Zeil, J. Pleiss. The scale-free nature of protein sequence space. *PLoS ONE* 13 (8): e0200815.

[6] a) D. J. Glover, D. Xu, D. S. Clark. Shaping the future of protein engineering. *Biochemistry,* **2019**, *58*, 1019-1021. b) D. Repecka, V. Jauniskis, L. Karpus, E. Rembeza, J. Zrimec, S. Poviloniene, I. Rokaitis, A. Laurynenas, W. Abuajwa, O. Savolainen, R. Meskys, M. K. M. Engqvist, A. Zelezniak. Expanding functional protein sequence space using generative adversarial networks. *bioRxiv,* **2019**. https://doi.org/10.1101/789719

[7] https://deepmind.com/research/case-studies/alphafold (10.12.2020)



2

Separating Thermodynamics from Kinetics – A New Understanding of the Transketolase Reaction

Transketolase catalyses asymmetric C-C bond formation of two highly polar compounds. Over the last 30 years, the reaction has unanimously been described in literature as irreversible due to the concomitant release of CO_2 when lithium hydroxypyruvate (LiHPA) is used as substrate. The reaction was now followed over an extended period of time, which revealed to be initially kinetically controlled. For the non-natural conversion of synthetically more interesting apolar substrates, a complete change of active site polarity is counterintuitively not necessary. Docking studies revealed water and hydrogen bond networks to be essential in substrate binding, thus allowing aliphatic aldehydes to be converted in the charged active site of transketolase.

This chapter is based on

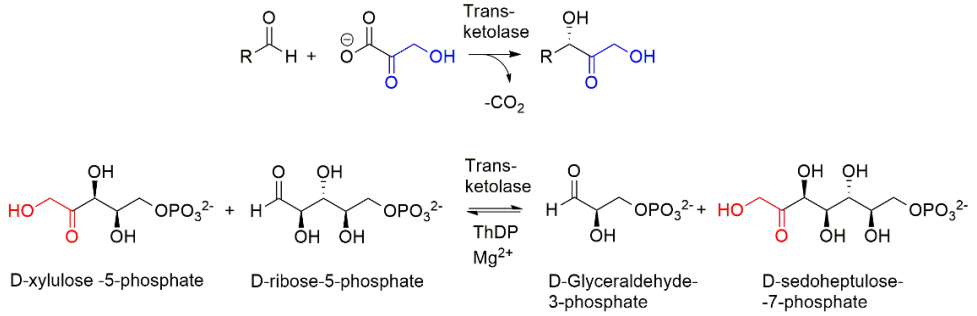
Stefan. R. Marsden, Lorina Gjonaj, Stephen J. Eustace and Ulf Hanefeld.

ChemCatChem, **2017**, *9*, 1808-1814.

DOI: 10.1002/cctc.201601649.

2.1 Introduction

Transketolase (TK, E.C. 2.2.1.1) is a Mg^{2+} and thiamine diphosphate (ThDP) dependent enzyme, which naturally catalyses the conversion of glycolysis derived metabolites into carbohydrates utilized for nucleotide synthesis and the production of essential aromatic amino acids *via* the Shikimate pathway.[1] The overall reaction comprises of the reversible transfer of a C_2-ketol group through asymmetric C-C bond formation. This makes the reaction interesting for synthetic applications. A multitude of enzymatic strategies have been developed in order to address the substantial importance of asymmetric C-C bond formation in organic synthesis, of which many rely on decarboxylation as driving force for the C_2-ketol transfer.[2-5]



Scheme 1: Decarboxylation driven use of hydroxypyruvate (HPA) as ketol donor for synthetic applications (top). Reversible, natural transketolase reaction (bottom).

Hydroxypyruvate (HPA) is currently used as the ketol donor of choice, as the Free Energy change from decarboxylation results in an equilibrium constant that entirely favours the product. For this reason, decarboxylation driven reactions are traditionally described as irreversible.[2-10] Indeed, the first *S. cerevisiae* TK (*Sc*TK) catalysed synthesis of L-erythrulose was performed with LiHPA to ensure it to be irreversible.[11-13] Yet in 2004, the coupling of two molecules of glycolaldehyde to L-erythrulose was also reported.[14] In

combination with the reversibility of the natural TK catalysed reactions, this renders an irreversible product formation unlikely from a mechanistic point of view. In recognition of the extensive use of decarboxylation in contemporary C-C bond formation strategies, a correct understanding of the actual impact of decarboxylation on the overall reaction is thus of great importance. In addition, decarboxylation has the disadvantage of a poor atom economy.

With phosphorylated polyols as typical substrates, TKs are naturally not disposed towards the conversion of aliphatic substrates. Nevertheless, *E. coli* TK has successfully been engineered by single-point mutations to convert a variety of aromatic and aliphatic aldehydes.[6,7] This is surprising, since the mutations introduced in *E. coli* TK do not render the active site more lipophilic.[6] *S. cerevisiae* TK shares 47% sequence identity with *E. coli* TK and the aligned crystal structures (1QGD and 1TRK) have an RMSD of 0.81 indicating extensive structural homology. Due to its facile heterologous overexpression in *E. coli*, *S. cerevisiae* TK was chosen as model enzyme to representatively investigate both the impact of decarboxylation in asymmetric C-C bond synthesis and the cause of enhanced activity towards aliphatic aldehydes that was previously observed for single-point mutations.[6,7]

2.2 Results and Discussion

Previous studies with *E. coli* TK mutants D469E and D469T showed, that highly polar or even charged amino acids could improve enzyme activity towards aliphatic aldehydes.[6] This is in contrast to other results, where non-phosphorylated substrates were better converted by TK mutants of reduced polarity (R528K, R528Q, R528K/S527T and R528Q/S527T).[15-16]

The equivalent mutations D477E and D477T were therefore created in *S. cerevisiae* TK to allow for a direct comparison.

Preparative scale reactions were in line with those reported for *E. coli* TK (Table 1). Again, mutant D477E was identified as the best variant for the conversion of aliphatic aldehydes.

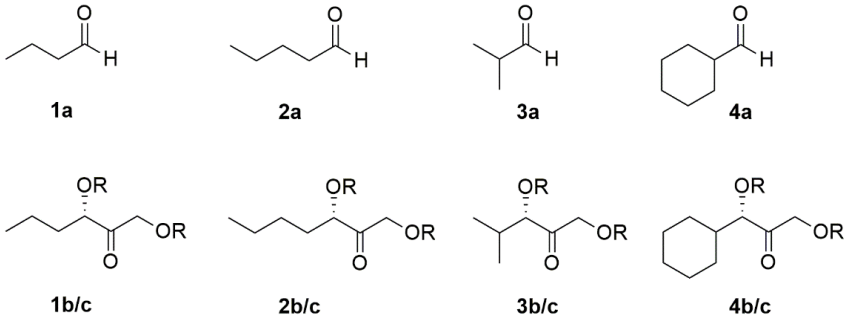


Figure 1: Overview of aldehyde substrates (entry **a**), (*3S*)-configured dihydroxyketone products (R = OH, entry **b**) and dibenzoylated products to allow for chiral separation (R = Bz, entry **c**).

Table 1: Isolated product yields and enantiomeric excess (*ee* %) of the (*S*)-configured major enantiomer. Conditions: 20 U of *S. cerevisiae* TK, 5 mM ThDP, 18 mM $MgCl_2$, 1 mmol LiHPA, 1 mmol aldehyde in 10 mL 5 mM NaPi, pH 7.0, 25°C, 200 rpm, 18 h.

	WT	D477E	D477T	R528K	R528Q	R528K/ S527T	R528Q/ S527T
1b	11±8% (*84%*)	34±15% (*94%*)	8% (*n.d.*)	10±8% (*81%*)	8±2% (*77%*)	8±3% (*73%*)	6±4% (*66%*)
2b	7% (*91%*)	61±13% (*90%*)	12±4% (*84%*)	6±4% (*82%*)	5±1% (*87%*)	6±1% (*68%*)	5±1 (*82%*)
3b	0% (*n.d.*)	41±20% (*99%*)	n.d. (*n.d.*)	3±1% (*n.d.)*	0% (*n.d.*)	0% (*n.d.*)	0% (*n.d.*)
4b	0% (*n.d.*)	0% (*n.d.*)	n.d. (*n.d.*)	0% (*n.d.*)	0% (*n.d.*)	0% (*n.d.*)	0% (*n.d.*)

Analysis of the Michaelis-Menten parameters further confirmed the results that were obtained on preparative scale (Table 2). Mutant D477E displayed the highest catalytic efficiency for the conversion of aliphatic aldehydes **1a** and **2a**, showing a 50 to 100-fold enhanced activity in comparison to the wild-type. While mutations at position R528, which natively binds to the phosphate group in phosphorylated substrates[15-16] and the incorporation of a group mutation strategy[17] did enhance enzyme activity, the improvements were only minor compared to the effect of mutation D477E.

Table 2: Michaelis-Menten parameters. Conditions: 50 µg purified *S. cerevisiae* TK, 1 mM ThDP, 4 mM $MgCl_2$, 100 mM LiHPA, 5-150 mM aldehyde, 5 mM NaPi, pH 7.0 25°C, 500 rpm. k_{cat} (s^{-1}), K_M (mM), k_{cat}/KM ($M^{-1}s^{-1}$).

		WT	D477E	D477T	R528K
1b	k_{cat}	1.2	42	0.5	0.8
	K_M	272	163	48	181
	$k_{cat}K_M^{-1}$	4.2	260	10	4.4
2b	k_{cat}	0.8	9.3	0.4	0.1
	K_M	327	40	43	16
	$k_{cat}K_M^{-1}$	2.4	233	9.9	6.9
3b	k_{cat}	0.4	0.6	n.d.	n.d.
	K_M	150	66		
	$k_{cat}K_M^{-1}$	2.9	8.3		

In silico docking studies

With a 50-100 fold improved catalytic efficiency towards **1a** and **2a** by variant D477E, we set out to investigate the resulting changes in the active site. The mutants were created *in silico* from the corresponding WT crystal structure (1GPU)[18] and the obtained model was energy minimised before docking of substrates **1a**-**4a** into the active site using YASARA.[19] The model

showed that extension of the carbon chain by mutating aspartate to glutamate (D477E) newly enabled hydrogen bond interactions between the glutamate carboxylate and the substrate carbonyl groups, bridged by a molecule of coordinated water at 1.7 Å each. In this manner, the substrate is correctly aligned towards the cofactor and the forming oxyanion is stabilised by charge delocalisation during nucleophilic attack. This interaction was exclusively predicted for the converted substrates **1a**-**3a** and not for the unconverted substrate **4a** (Figure 2). In agreement with the kinetic data, these docking studies illustrate, that correct substrate orientation towards the activated cofactor (improving not only k_{cat}, but potentially also K_M) is of greater importance for catalysis than an increase based solely on substrate affinity (improving only K_M). This would also explain why the introduction of an isoleucine residue at the equivalent position in transketolase from *Geobacillus stearothermophilus* did not lead to large rate improvements.[20]

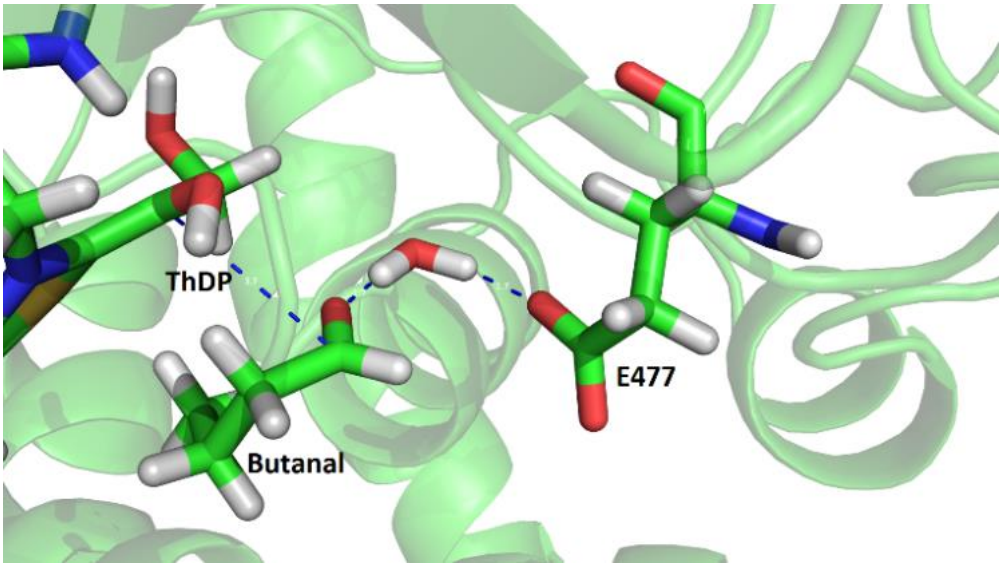
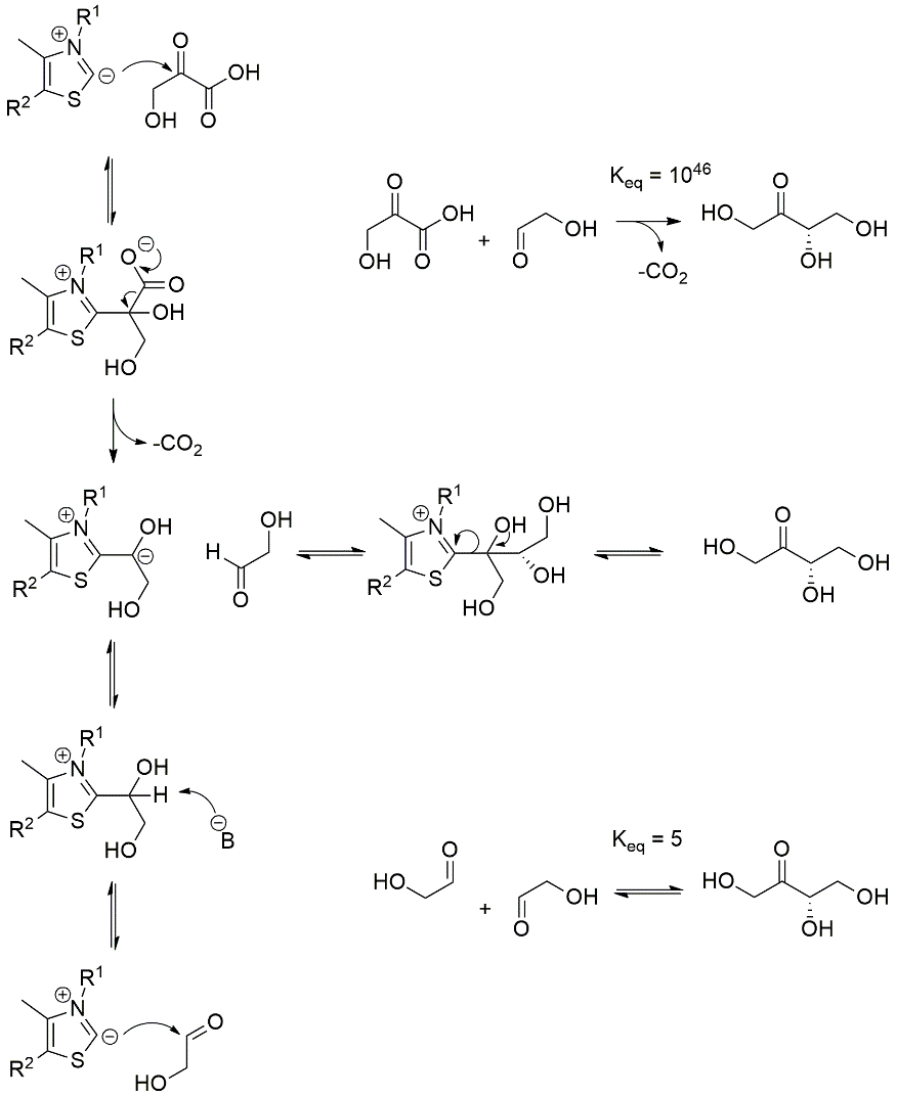


Figure 2: *In silico* docking of butanal into the energy-minimised active site of D477E TK using YASARA, showing a hydrogen bond network *via* a bridging water molecule.

Mechanistic analyses

The change in Gibb's Free Energy (Δ_rG^0) for the synthesis of L-erythrulose from glycolaldehyde and LiHPA amounts to -264.5 kJ/mol under standard conditions, largely due to the contribution of decarboxylation. This corresponds to an equilibrium constant of $K_{eq} = 10^{46}$ in favour of the product. In contrast, the one-substrate reaction couples two molecules of glycolaldehyde to L-erythrulose [14] and shows an equilibrium constant of $K_{eq} = 5.0$ ($\Delta_rG^0 = -4.0$ kJ/mol). Supported by the mechanistic reversibility of the natural reaction, the one-substrate reaction should therefore be a true equilibrium reaction. In the proposed mechanism for TK catalysed reactions with lithium hydroxypyruvate, the thermodynamically irreversible decarboxylation of LiHPA effects the direct formation of the carbanion on the activated ketol. For the one-substrate reaction however, the activated carbanion must be formed by catalytic deprotonation from residue His481 as alternative to decarboxylation, generating the activated intermediate at a lower rate in comparison to its generation by decarboxylation. At the stage of the activated ketol bearing the carbanion, the enzyme can no longer distinguish whether it was formed *via* a reaction pathway involving decarboxylation, or *via* catalytic deprotonation. The information about the thermodynamic driving force of decarboxylation is therefore already lost prior to the actual product formation. This mechanism consequently suggests that TK catalysed synthesis reactions are reversible *via* the mechanism of the one-substrate reaction, which would split the product back into the corresponding aldehyde and one molecule of glycolaldehyde. The thermodynamic contribution of decarboxylation should therefore not affect the position of the overall equilibrium (Scheme 3) and argue against an irreversible product formation. In conclusion, it should thus be possible to avoid the release of CO_2 and to improve the atom economy.



Scheme 3: Catalytic mechanism of the decarboxylation driven transketolase reaction (top) and the one-substrate reaction *via* catalytic deprotonation (bottom).

Equilibrium analysis

In order to experimentally confirm the reversibility of TK catalysed reactions, L-erythrulose was synthesized both *via* the one-substrate reaction and in the decarboxylation driven fashion with LiHPA. The reactions were performed in sealed NMR tubes, allowing for direct measurements of the product L-erythrulose.[21] LiHPA was completely converted within 30 minutes, while the complex oligomerization and hydration of glycolaldehyde in aqueous solution did not allow for its straightforward analysis.[22] Both reactions were quantitatively followed over an extended period of time.

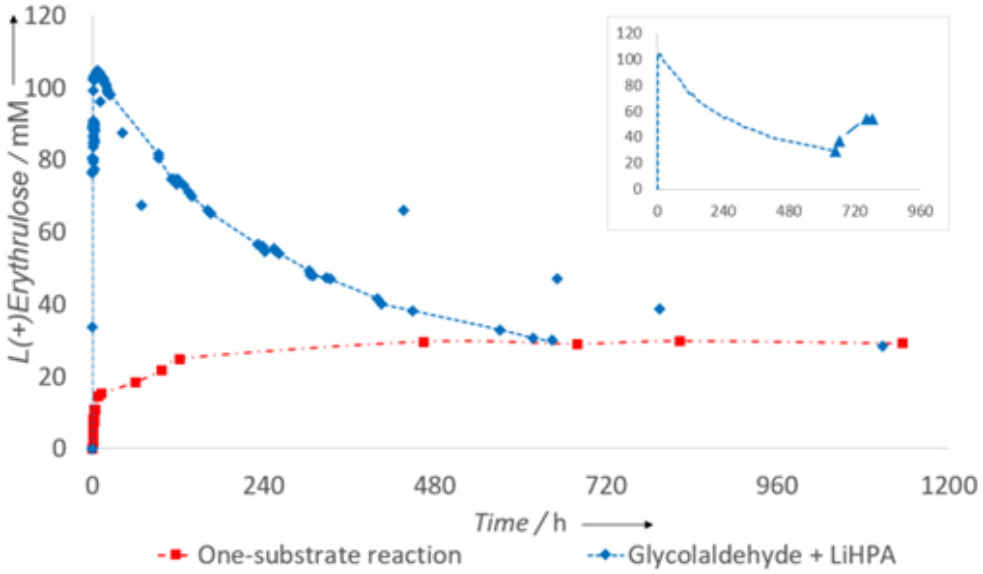


Figure 3: Kinetically controlled synthesis of L-erythrulose with LiHPA (blue) and thermodynamically controlled one-substrate reaction (red). Addition of fresh LiHPA showed, that the enzyme had retained its activity over the whole time course (inset).

The one-substrate reaction was found to be limited to less than 30% conversion by the position of the thermodynamic equilibrium. In contrast, complete conversion was swiftly observed with LiHPA as substrate. If product formation was irreversible due to the release of CO_2, the reaction should stop here. However, in line with a mechanistically reversible reaction, a slow decline of L-erythrulose concentration was subsequently observed, converging towards the equilibrium concentration of the one-substrate reaction at K_{eq} = 29.1 ± 0.6 mM. The synthesis reaction was thus shown to benefit from a kinetic effect, enabling high yields at the beginning of the reaction. The mechanism of the reverse reaction subsequently effected the thermodynamic equilibration over a time course of several weeks. In order to confirm that the observed equilibration indeed was enzyme catalysed, another portion of LiHPA was added at the end. Retained enzymatic activity was observed, while control reactions without enzyme showed no conversion.

The representative formation of L-erythrulose from glycolaldehyde and LiHPA was thus shown to be initially kinetically controlled, contrary to earlier assumptions about the thermodynamic driving force of CO_2 release. The proposed reaction mechanisms depicted in scheme 3 suggests these findings to generally hold true for all TK catalysed reactions with HPA. Following the example of the pyruvate decarboxylase catalysed synthesis of (*R*)-phenylacetylcarbinol with acetaldehyde replacing the traditional donor substrate pyruvate,[23] the development of novel strategies which do not rely on decarboxylation is of commercial relevance. To do so, a correct understanding of decarboxylation is of utmost importance. In syntheses where aldehydes other than glycolaldehyde are used as acceptors, formation of the desired product will be competing with the one-substrate reaction. Active site engineering as pioneered by Pohl for a range of ThDP dependent enzymes could ensure that glycolaldehyde will be used as the sole donor molecule in mixed carboligation reactions.[24]

2.3 Conclusion

In this study, the creation of a new hydrogen bond network was shown to be an effective strategy for increasing a substrate's affinity towards the active site. The introduction of highly polar, hydrogen bond donating residues (Asp, Glu, Ser, Thr, Tyr, Lys, Arg) may therefore benefit the conversion apolar substrates, whereas a polarity-based analysis would suggest the opposite. This alternative approach for the rational mutagenesis of TKs towards hydrophobic substrates was demonstrated. While decarboxylation driven C-C bond formation reactions traditionally are misinterpreted in literature as irreversible, mechanistic analyses and experimental evidence unambiguously showed that the reaction is under kinetic control. In the context of man-made climate change, we thus have to re-evaluate our choice of donor substrates and the use of decarboxylation strategies in synthetic applications.

Acknowledgements

Financial support from STW (grant 11142) to L. G. is gratefully acknowledged.

2.4 Methods Summary

Detailed experimental protocols next to additional graphs are available in the supplementary information of the *open access* online version of this article.

Preparative scale reactions were conducted with crude cell free extracts (20U TK activity based on glycolaldehyde activity assay) in sodium phosphate buffer (5 mM NaPi, 18 mM ThDP, 5 mM Mg^{2+}, pH = 7.0). LiHPA (110 mg, 1.0 mmol, 1.0 eq.) and the corresponding aldehyde (1.0 mmol, 1.0 eq.) were added to a final volume of 10 mL. Reactions were carried out overnight in sealed flasks (25°C, 200 rpm). The product was extracted with MTBE and the solvent was removed *in vacuo*.

Chiral analysis to determine the enantiomeric excess was performed by HPLC upon derivatisation with benzoyl chloride on a Chiralpak AD-H column (0.46x25 cm, Daicel, *n*-heptane/*i*-PrOH 97:3, 35°C, 1 mL/min, 280 nm). The dihydroxyketone substrate (1.0 eq.) was dissolved in dry dichloromethane (10 mL) and dry triethylamine (10.0 eq.) and benzoyl chloride (10.0 eq.) were added under inert conditions. The reaction

mixture was stirred for two hours at room temperature before it was quenched by the addition of sat. $NaHCO_3$. The organic phase was washed with sat. $NaHCO_3$, sat. NH_4Cl and brine. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (petrolether / EtOAc 10:1). Racemic standards were synthesized from LiHPA and the corresponding aldehyde using the previously described biomimetic transketolase reaction with *N*-methylmorpholine.[31]

Glycolaldehyde activity assay.[15] The volumetric activity of cell free extracts was determined by incubating 50 µL with LiHPA (50 mM) and glycolaldehyde (50 mM) in 300 µL total reaction volume (5 mM NaPi, 5 mM ThDP, 18 mM Mg^{2+}, 25°C, 800 rpm, 15 min, 1.5 mL Eppendorf tubes). The reaction mixture was quenched by addition of TFA (300 µL, 0.2% v/v), centrifuged and analysed by HPLC on an IC-Sep Coregel 87H3 column (0.4x25 cm, Transgenomic, 0.1% (v/v) TFA, pH = 2.5, 60°C, 0.8 mL/min, 210 nm) using a calibration curve as an external standard.

Michaelis-Menten parameters were determined under initial rate conditions (<20% conversion) by adjusting the reaction time accordingly. LiHPA (100 mM) and varying concentrations of the corresponding aldehyde (5 – 150 mM) were incubated with holo-transketolase (50 µg/337 pmol, 1 mM ThDP, 4 mM Mg^{2+}, 25°C, 500 rpm, n = 2). The reactions were quenched by the addition of TFA (1:1, 0.2% v/v), centrifuged and analysed by HPLC on an IC-Sep Coregel 87H3 column (0.4x25 cm, Transgenomic, 0.1% (v/v) TFA, pH = 2.5, 60°C, 0.8 mL/min, 210 nm) using a calibration curve as an external standard.

Equilibrium analysis by NMR.[21,32] Holo-*Sc*TK (200 µg, 1.35 nmol, 5 mM ThDP, 18 mM Mg^{2+}, 5 mM NaPi, pH 7.0, 25°C, 500 µL, sealed NMR tube) was incubated with glycolaldehyde (100 mM) and LiHPA (100 mM) for the decarboxylation driven reaction, or 200 mM glycolaldehyde in the case of the one-substrate reaction respectively. NMR spectrums were recorded using PRESAT for water suppression and the benzene signal from an NMR capillary was used as external reference for quantitation. Formation of L-erythrulose was followed *via* its characteristic peaks 4.61 (1H, d, $^2J_{HH}$ 19.6 Hz), 4.52 (1H, d, $^2J_{HH}$ 19.6 Hz).

Computational docking studies. *In silico* docking studies were carried out with YASARA (Version 16.2.18) using the crystal structures 1TRK (free ThDP cofactor) and 1GPU (containing the covalent ketol intermediate) for *S. cerevisiae* TK and 1QGD for *E. coli* TK. The simulation box was defined at 10 Å around the thiamine C2 in 1TRK and around the ylid anion in 1GPU. The substrates were energy minimized with ChemBio3D Ultra 12.0 (Cambridgesoft) using MM2 energy minimizations. The mutation D477E was introduced into 1GPU and the model was subsequently energy minimized using YASARA before docking.

References

[1] a) S. Hohmann, P. A. Meacock, *Biochim. Biophys. Acta*, **1998***,1385*, 201-219; b) R. Wohlgemuth, *J. Mol. Catal. B: Enz*., **2009**, *61*, 23-29; c) R. Kluger, K. Tittmann, *Chem. Rev*. **2008**, *108*, 1797-1833.

[2] M. Pohl, C. Wechsler, M. Müller in *Science of Synthesis, Biocatalysis in Organic Synthesis 2,* (Eds.: K. Faber, W.-D. Fessner, N. J. Turner), Georg Thieme Verlag, Stuttgart, **2014**, pp. 93-127.

[3] M. Brovetto, D. Gamenara, P. S. Méndez, G. A. Seoane, *Chem. Rev.*, **2011**, *111*, 4346-4403.

[4] J. Sukumaran, U. Hanefeld, *Chem. Soc. Rev.*, **2005**, *34*, 530-542.

[5] A. Ranoux, U. Hanefeld in *Stereoselective Synthesis of Drugs and Natural Products*, (Eds.: V. Andrushko, N. Andrushko), John Wiley & Sons, Inc., **2013**, pp. 831-857.

[6] A. Cázares, J. L. Galman, L. G. Crago, M. E. B. Smith, J. Strafford, L. Ríos-Solís, G. J. Lye, P. A. Dalby, H. C. Hailes, *Org. Biomol. Chem.,* **2010**, *8*, 1301-1309.

[7] J. L. Galman, D. Steadman, S. Bacon, P. Morris, M. E. B. Smith, J. M. Ward, P. A. Dalby, H. C. Hailes, *Chem. Commun.,* **2010**, *46*, 7608-7610.

[8] F. Subrizi, M. Cárdenas-Fernández, G. J. Lye, J. M. Ward, P. A. Dalby, T. D. Sheppard, H. C. Hailes, *Green Chem.*, **2016**, *18*, 3158-3165.

[9] T. Saravanan, M. Reif, D. Yi, M. Lorillière, F. Charmantray, L. Hecquet, W.-D. Fessner, *Green Chem.*, **2017**, *19*, 481-489.

[10] M. Lorillière, M. De Sousa, F. Bruna, E. Heuson, T. Gefflaut, V. de Berardinis, T. Saravanan, D. Yi, W.-D. Fessner, F. Charmantray, L. Hecquet, *Green Chem.*, **2017**, *19,* 425-435.

[11] U. Schörken, G. A. Sprenger, Biochim*. Biophys. Acta,* **1998**, *1385*, 229-243.

[12] Y. Kobori, D. C. Myles, G. M. Whitesides, *J. Org. Chem.*, **1992**, 5899-5907.

[13] J. Bolte, C. Demuynck, H. Samaki, *Tetrahedron Lett.*, **1987**, *28*, 45, 5525-5528.

[14] I. A. Sevostyanova, O. N. Solovjeva, G. A. Kochetov, *Biochem. Biophys. Res. Commun.*, **2004**, *313*, 771-774.

[15] A. Ranoux, S. K. Karmee, J. Jin, A. Bhaduri, A. Caiazzo, I. W. C. E. Arends, U. Hanefeld, *ChemBioChem,* **2012**, *13*, 1921-1931.

[16] A. Ranoux, U. Hanefeld, *Top. Catal.*, **2013**, *56*, 750-764.

[17] a) M. T. Reetz, M. Bocola, J. D. Carballeira, D. Zha, A. Vogel, *Angew. Chem. Int. Ed.*, **2005**, *44*, 4192-4196; b) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature*, **2012**, *485*, 185-194.

[18] E. Fiedler, S. Thorell, T. Sandalova, R. Golbik, S. König, G. Schneider, *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 591-595.

[19] E. Krieger, G. Vriend, *Bioinformatics,* **2014**, *30*, 2981-2982.

[20] D. Yi, T. Saravanan, T. Devamani, F. Charmantray, L. Hecquet, W.-D. Fessner, *Chem. Commun*., **2015**, *51*, 480-483.

[21] T. Purkarthofer, W. Skranc, H. Weber, H. Griengl, M. Wubbolts, G. Scholz, P. Pöchlauer, *Tetrahedron*, **2004**, *60*, 735-739.
[22] J. Kua, M. M. Galloway, K. D. Millage, J. E. Avila, D. O. de Haan, *J. Phys. Chem. A*, **2013**, *117*, 2997-3008.
[23] B. Rosche, M. Breuer, B. Hauer, P. L. Rogers, *Biotechnol. Bioeng.,* **2004**, *86*, 788-794.
[24] R. Westphal, C. Vogel, C. Schmitz, J. Pleiss, M. Müller, M. Pohl, *Angew. Chem. Int. Ed.*, **2014**, *53*, 9376-9379.
[25] P. Hendil-Forssell, M. Martinelle, P.-O. Syrén, *Chem. Commun.*, **2015**, *51,* 17221-17224.
[26] C. Kürten, B. Carlberg, P.-O. Syrén, *Catalysts,* **2016**, *6,* 90.
[27] S. Matsuoka, S. Sugiyama, D. Matsuoka, M. Hirose, S. Lethu, H. Ano, T. Hara, O. Ichihara, S. Roy Kimura, S. Murakami, H. Ishida, E. Mizohata, T. Inoue, M. Murata, *Angew. Chem. Int. Ed.*, **2015**, *20*, 39, 1508-1511.
[28] P.-O. Syrén, S. C. Hammer, B. Claasen, B. Hauer, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4845-4849.
[29] B. Breiten, M. R. Lockett, W. Sherman, S. Fujita, M. Al-Sayah, H. Lange, C. M. Bowers, A. Heroux, G. Krilov, G. M. Whitesides, *J. Am. Chem. Soc.*, **2013**, *135*, 15579-15584.
[30] K. G. Morris, M. E. B. Smith, N. J. Turner, *Tetrahedron: Asymmetry*, **1996**, *7*, 8, 2185-2188.
[31] M. E. B. Smith, K. Smithies, T. Senussi, P. A. Dalby, H. C. Hailes, *Eur. J. Org. Chem.*, **2006**, *5*, 1121-1123.
[32] M. Weber, C. Hellriegel, A. Rueck, J. Wüthrich, P. Jenks, *J. Pharm. Biomed. Anal.*, **2014**, *93*, 102-110.

3

Assessing the ThDP-dependent *Ec*PDH E1 subunit for carboligation reactions with aliphatic ketoacids

The synthetic properties of the ThDP-dependent pyruvate dehydrogenase E1 subunit from *Escherichia coli* (*Ec*PDH E1) was assessed for carboligation reactions with aliphatic ketoacids. Due to its role for metabolism, *Ec*PDH E1 was previously characterised with respect to its biochemical properties, but it was never applied for synthetic purposes. Here we show, that *Ec*PDH E1 is a promising biocatalyst for the production of chiral α-hydroxyketones. WT *Ec*PDH E1 shows a 180-250 fold higher catalytic efficiency towards 2-oxobutyrate or pyruvate, respectively, in comparison to engineered transketolase variants from *Geobacillus stearothermophilus* (TK_{GST}). Its broad active site cleft allows for the efficient conversion of both (*R*)- and (*S*)-configured α-hydroxyaldehydes, next to linear and branched aliphatic aldehydes as acceptor substrates under kinetically controlled conditions. The alternate, thermodynamically controlled self-reaction of aliphatic aldehydes was shown to be limited to low levels of conversion, which we propose to be due to their large hydration constants. Additionally, the thermodynamically controlled approach was demonstrated to suffer from a loss of stereoselectivity, which makes it unfeasible for aliphatic substrates.

This chapter is based on

Stefan R. Marsden, Duncan G. G. McMillan and Ulf Hanefeld

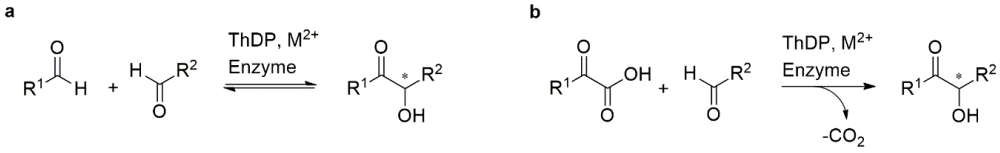
International Journal of Molecular Sciences, **2020**, *21*, (22), 8641.

DOI: 10.3390/ijms21228641

3.1 Introduction

Thiamine diphosphate (ThDP) dependent enzymes are excellent biocatalysts for the synthesis of chiral α-hydroxyketones (acyloins) from two aldehydes.[1,2] The reaction is initiated by the activated ThDP cofactor forming a covalent intermediate with the donor substrate. This induces an 'Umpolung' that turns the carbonyl group into a nucleophile.[3,4] In spite of being 100% atom efficient, unfavourable equilibrium conditions can limit this thermodynamically controlled approach to low levels of conversion.[5]

This issue can be addressed *via* the decarboxylation of ketoacids as donor substrate analogues, which renders the reaction kinetically controlled and allows for complete conversion (Scheme 1).[5,6] This feature makes the application of ThDP-dependent enzymes interesting for industrial applications.[7] Their acceptor substrate scope was engineered to allow for the conversion of aliphatic-,[5,8,9] aromatic-[10] and non-phosphorylated[11,12] aldehyde substrates with enhanced or reversed stereoselectivity,[13-15] of which comprehensive reviews were published elsewhere.[16-18] Yet, the expansion of the donor substrate scope by mutagenesis remains a formidable challenge.



Scheme 1: (a) Thermodynamically controlled conversion of aldehyde substrates. A lack of discrimination between the donor and acceptor substrate can lead to additional regioisomers, which typically limits this approach to $R^1 = R^2$. **(b)** Decarboxylation renders the reaction kinetically controlled and allows for mixed carboligation reactions ($R^1 \neq R^2$) in a controlled fashion. Monitoring of the reaction's progress is required to prevent subsequent thermodynamic equilibration towards the corresponding aldehyde starting materials.[5,6]

A multitude of essential amino acid interactions are required for the binding and activation of the ThDP cofactor and must not be disrupted if function is to be retained.[19] A metal binding site coordinates a divalent cation (typically Mg^{2+}), which ionically binds the ThDP cofactor *via* its pyrophosphate group. A distal glutamate and histidine residue then function together as catalytic base to convert the cofactor into its active ylide state *via* deprotonation of the thiazole ring.[20,21] To spatially allow for the required proton transfer, the ThDP cofactor adopts an energetically disfavoured V-conformation that is evolutionary conserved within ThDP-dependent enzymes.[22-24] Additional interactions with the covalent intermediate further promote the distorted conformation and prevent its relaxation into a lower energy state. The activation energy for the following C-C bond formation is thereby reduced.[25] The holoenzyme is usually a homodimer with two symmetrical active sites located at the dimer interface.

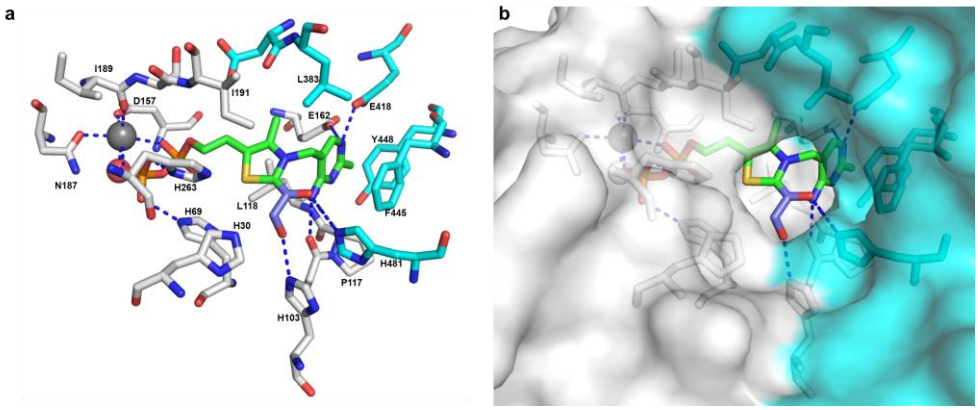


Figure 1: (a) Interactions between the ThDP cofactor (green) and residues from both monomers (white/cyan) in *Saccharomyces cerevisiae* transketolase (*Sc*TK). The covalent C_2-ketol intermediate is stabilised by hydrogen bonds with two histidine residues. **(b)** Surface view of the active site in WT *Sc*TK. The ThDP cofactor is buried within the enzyme, and only accessible *via* a narrow substrate channel. This sterically limits the size of the donor substrate to the transfer of a C_2-ketol unit in WT *Sc*TK. Dotted lines denote H-bonds or ionic interactions at 2-4 Å distance. Figures were created in PyMol from 1gpu.pdb.

A proton wire allows for communication between the two active sites, making them non-equivalent in terms of cofactor affinity and effects an alternating half-of-the-sites reactivity.[32,33,36,37] A switch from positive, cooperative binding of ThDP ($n \approx 2$ at 1 mM Mg^{2+}) to negative cooperativity ($n = 0.61$ at 3 mM of Mg^{2+}) highlights the complexity of interactions in *E. coli* TK.[34] Furthermore, oxidative stress during protein expression was shown to result in the post-translational oxidation of Cys157 to a sulfenic acid group in *E.coli* TK. This oxidation leads to a 100-fold increased affinity towards ThDP and a 20-fold increase in transketolase activity.[34]

Examples for the conversion of non-natural donor substrates by ThDP-dependent enzymes are scarce, due to the highly specific interactions within their active sites (Figure 1). This feature constitutes a major drawback in terms of engineering, as it limits the structural diversity that is currently accessible with ThDP dependent enzymes. Examples include the decarboxylative conversion of (*S*)-4-hydroxy-2-oxoglutarate with the enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) from *E. coli* using a kinetically controlled approach,[38] while thermodynamically controlled one-substrate, benzoin-type reactions were demonstrated with benzaldehyde lyase (BAL) from *Pseudomonas fluorescens biovar I* using benzaldehyde derivatives.[39]

Notably, a thermostable transketolase from *Geobacillus stearothermophilus* (TK_{GST}) was successfully engineered by employing successive rounds of iterative site-saturation mutagenesis towards the conversion of pyruvate, 2-oxobutyrate and 3-methyl-2-oxobutyrate as donor substrates,[40] of which triple variants were recently reported in a follow-up study.[41] Similarly, *E. coli* TK was also engineered for the conversion of pyruvate.[42]

However, transketolases naturally evolved to specifically accept highly polar, phosphorylated carbohydrates as natural substrates,[43] which makes their wild-type variants inept for the conversion of aliphatic ketoacids (e.g. pyruvate). For this reason, engineered TK variants still tend to display rather low catalytic efficiencies towards these non-natural substrates.[40-42] Since ThDP dependent enzymes

comprise of a modular structure and share the same catalytic mechanism,[44] we hypothesised that the screening of different enzyme scaffolds could prove more suitable for the conversion of aliphatic ketoacids in carboligation reactions.

The Thiamine Enzyme Engineering Database (TEED) provides an excellent overview of the nine superfamilies of ThDP-dependent enzymes, including sequence and structural data.[45-47] During a qualitative database search, the ThDP-dependent pyruvate dehydrogenase E1 subunit from *E. coli* (*Ec*PDH E1) attracted our attention. As part of the pyruvate dehydrogenase complex (PDHc), the E1 subunit catalyses the decarboxylative activation of pyruvate for the synthesis of acetyl-CoA (Figure 2a).[48] Due to its relevance for metabolism, E1 was extensively characterised with respect to its biochemical properties, but has never been applied for synthetic purposes.[49]

*Ec*PDH E1 does not require the presence of the E2 and E3 subunits and purified *Ec*PDH E1 is fully active alone.[50] With pyruvate as its natural donor substrate, *Ec*PDH E1 evolved towards the efficient conversion of aliphatic ketoacids and shows an inherent promiscuity towards 2-oxobutyrate.[51] The use of a sterically demanding lipoate moiety as its natural acceptor substrate requires a wide substrate channel, and should therefore enable the conversion of a broad range of substrates (Figure 2b). This is in stark contrast to the narrow substrate channel of transketolases (Figure 1b). Notably, this also implies that *Ec*PDH E1 should not display a requirement for phosphorylated substrates; a feature which commonly impairs the catalytic efficiency of transketolases.[11] These properties motivated us to assess the synthetic performance of *Ec*PDH E1 for the conversion of aliphatic ketoacids in carboligation reactions.

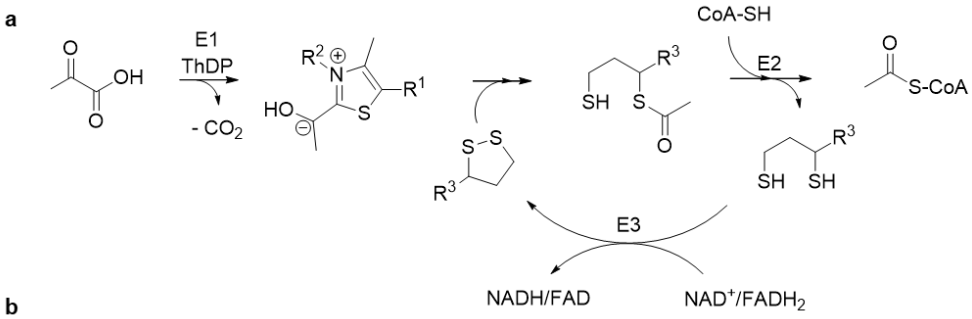


Figure 2: (a) Natural reaction of the pyruvate dehydrogenase complex (PDHc): ThDP-dependent E1 forms a covalent intermediate with pyruvate and catalyses its decarboxylation. Subsequent C-S bond formation with a bulky lipoate moiety affords a thioester. Transesterification by E2 then affords acetyl-CoA, and Flavin-mediated oxidation by E3 closes the catalytic cycle. **(b)** Surface view of the active site in WT *Ec*PDH E1. A broad active site cleft allows for the accommodation of sterically more demanding substrates. Figures were created with PyMol from 2iea.pdb

3.2 Results and discussion

Expression and purification

The aceE gene (encoding for *Ec*PDH E1, accession number P0AFG9, EC 1.2.4.1) was codon optimised for recombinant expression in *E. coli*. The target gene was subsequently cloned into the pBAD/HisA expression plasmid using the designed KpnI and XhoI restriction sites. This cloning strategy introduced an *N*-terminal His_6-tag, which is separated from the *N*-terminus by a linker of 32 amino acids. *E. coli* Top10 cells were transformed with the final construct, and the enzyme was expressed in a batch fermentation until an increase in dissolved oxygen indicated the depletion of nutrients.

*Ec*PDH E1 was subsequently purified by affinity chromatography *via* its *N*-terminal His_6-tag to give a pure protein yield of 350 mg/L of expression medium under non-optimised conditions. Notably, the disruption of cells by ultrasonication on ice led to a complete loss of activity, while active enzyme was obtained with a cell disrupter using 3 passes at 1.8 kbar in combination with lysozyme.

Optimisation of reaction conditions

A spectrophotometric assay utilising 2,6-dichloroindophenol (DCPIP) and pyruvate as substrates was initially used to identify optimal conditions for *Ec*PDH E1 catalysed reactions.[52] The reduction of DCPIP was followed at its isosbestic point (517 nm) to allow for the determination of the optimal pH. E1 retained its activity over a broad pH range from pH 5.5 to 9.5, with its highest activity at pH 7.5 (Figure 3a). The enzyme showed good thermal stability with no loss of activity after two hours of incubation at 40°C. However, higher temperatures swiftly lead to its complete inactivation (Figure 3b). Notably, *Ec*PDH E1 can be stored with no loss of activity for at least six months at -20°C (5 mg/mL *Ec*PDH E1 in 20 mM potassium phosphate buffer, pH 7.0).

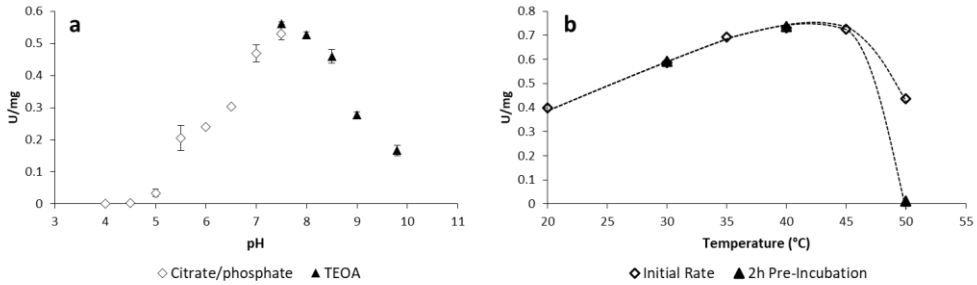


Figure 3: (a) A pH optimum of 7.5 was determined for *Ec*PDH E1 activity by the DCPIP assay (0.2 mM ThDP, 2 mM $MgCl_2$, 2mM sodium pyruvate, 100 µM DCPIP, 20 mM KPi, 25°C). **(b)** *Ec*PDH E1 displays a maximum activity at 40°C with no concomitant loss of activity after two hours. Higher temperatures swiftly reduce the initial rate and lead to a complete inactivation (0.2 mM ThDP, 2 mM $MgCl_2$, 2 mM sodium pyruvate, 100 µM DCPIP, 20 mM KPi, pH 7.5). Control reactions were performed in the absence of enzyme.

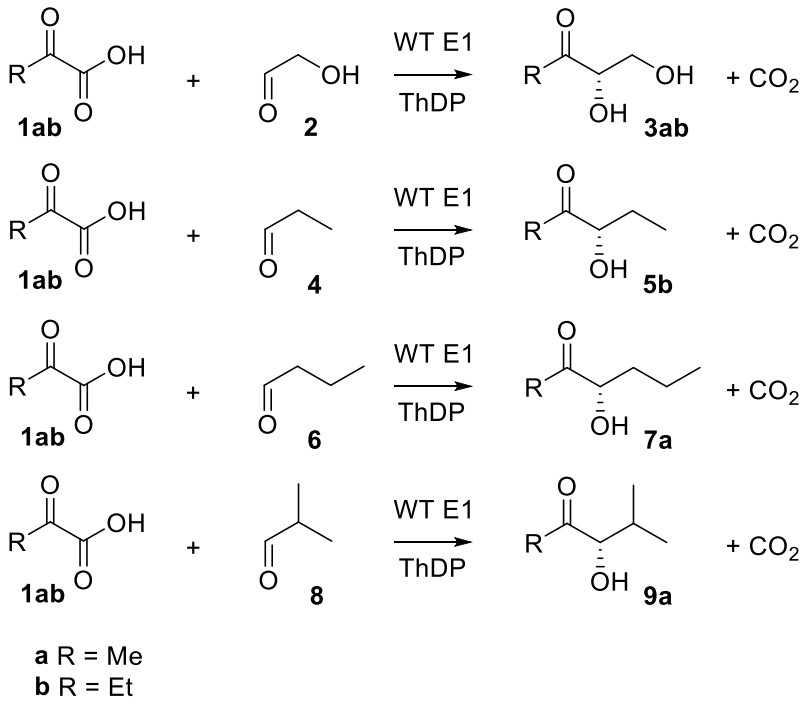
Preparative scale reactions

Having identified optimal conditions for *Ec*PDH E1 catalysed reactions, we set out to investigate its substrate scope on a preparative scale. A small excess of 1.2 equivalents of either the donor or the acceptor substrate was used where appropriate, in order to benefit the subsequent workup by extraction. Both pyruvate and 2-oxobutyrate were readily accepted as donor substrates.

Regarding the acceptor scope, polar α-hydroxyaldehydes were efficiently converted next to apolar, linear and branched aliphatic aldehydes (Scheme 2). *Ec*PDH E1 displayed (*S*)-selectivity, and the corresponding acyloins were consistently obtained in good enantiomeric purity (93 - 95% *ee*) and poor to good yields (5 - 70%, Table 1) due to their volatility. Similar enantiomeric purities were reported for engineered TK_{GST} variants using α-hydroxyaldehyde acceptor substrates.[40] However, the combination of 2-oxobutyrate as donor substrate with aliphatic aldehyde acceptors previously led to an almost complete loss of

stereoselectivity (6 - 33% *ee*) in TK_{GST} variants,[41] which can now be produced in good enantiomeric purity by *Ec*PDH E1.

The isolated products were subsequently used as external standards for a more detailed kinetic analysis of WT *Ec*PDH E1. by HPLC.



Scheme 2: Preparative scale (0.6-10 mmol) coupling of ketoacids **1a-b** with various aldehydes to afford the corresponding (*S*)-configured acyloins under kinetically controlled conditions.

Table 1: Synthetic performance of *Ec*PDH E1 on preparative scale. *Ec*PDH E1 consistently afforded the acyloin products in good enantiomeric purity from both aliphatic and hydroxyaldehyde substrates. Conditions: 0.5 mg/mL WT *Ec*PDH E1, 100-120 mM of each substrate loading, 20 mM KPi, pH 7.0, 24 hours, room temperature. Conversions were determined with respect to 1.0 equivalents of the ketoacid substrate.

Product	Conversion (%)	Isolated yield (%)	*ee* (%)
3a	64	5	93
3b	44	43	93
5b	78	68	95
7a	99	57	95
9a	69	22	93

Conversion of (*R*)- and (*S*)-configured α-hydroxyaldehydes

Transketolases catalyse the kinetic resolution of α-hydroxyaldehydes, displaying a strong stereopreference for the (*2R*)-configuration.[53,54] While this may be desirable in some cases, this feature prevents the conversion of (*2S*)-configured substrates and limits the product scope when enantiopure substrates are available.

To investigate a possible stereopreference regarding the configuration of α-hydroxyaldehyde substrates, the reaction between racemic DL-glyceraldehyde and pyruvate was followed over time. Complete conversion of the racemic substrate was readily achieved, and the time course did not indicate a notable discrimination between D- and L-glyceraldehyde. With lipoate as its natural acceptor substrate, *Ec*PDH E1 possesses a broad active site cleft that does not require residues for the recognition of hydroxyl groups. E1 therefore displays an extended product scope over transketolases by allowing for the efficient conversion of both (*R*)- and (*S*)-configured α-hydroxyaldehydes.

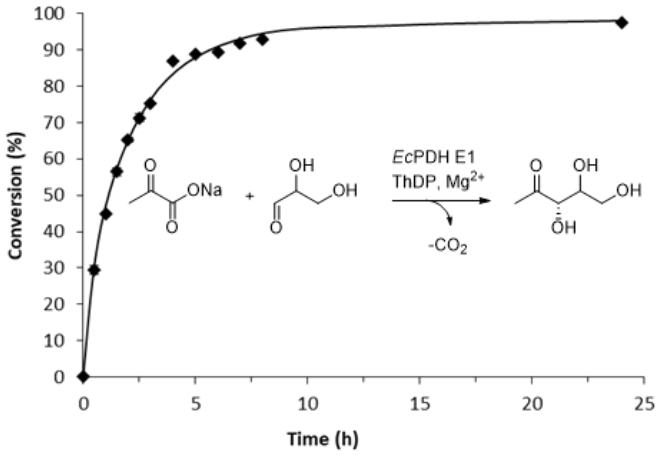


Figure 4: Time course of the *Ec*PDH E1 catalysed conversion of racemic DL-glyceraldehyde and sodium pyruvate (0.1 mg/mL *Ec*PDH E1, 0.2 mM ThDP, 2 mM $MgCl_2$, 50 mM sodium pyruvate, 50 mM DL-glyceraldehyde, 20 mM KPi, pH 7.5, n = 3). Complete conversion indicated no stereopreference regarding the configuration of α-hydroxyaldehyde substrates.

Thermodynamically controlled one-substrate reactions

Benzaldehyde lyase (BAL) catalyses the thermodynamically controlled coupling of aromatic aldehydes in benzoin-type reactions, which proceeds with 100% atom economy and was reported to afford the products in both high enantiomeric purity and yield.[39] Similarly, transketolase catalyses the self-reaction of glycolaldehyde to erythrulose,[55] albeit with low conversion due to an unfavourable equilibrium constant.[5]

With this in mind, the efficiency of *Ec*PDH E1 to catalyse the self-reaction of aliphatic aldehydes was explored. (*4S*)-hydroxyhexan-3-one **5b** can be synthesised from 2-oxobutyrate and propionaldehyde under kinetically controlled conditions, so its alternate synthesis *via* the self-reaction of propionaldehyde was also examined. The reaction was found to converge towards 10% conversion, and equilibrium conditions were demonstrated by the addition of extra enzyme; no change was observed (Figure 5). The examples of *Sc*TK and *Ec*PDH E1 catalysed

self-reactions are in stark contrast to the performance of BAL catalysed conversions with aromatic aldehydes, where high yields were reported.[39]

Notably, aldehydes can form hydrates in aqueous solution, and their hydration constant is determined by the electrophilicity of the carbonyl group. This hydration of aldehydes constitutes a competing side reaction, which influences the maximal extent of conversion. Aromatic aldehydes are stabilised through resonance, and the hydration of benzaldehyde is only minor. In contrast, propionaldehyde shows a 63-fold higher hydration constant, which further increases with the presence of electron withdrawing residues (e.g. a hydroxyl group in glycolaldehyde).[56]

The viability of thermodynamically controlled self-reactions of aldehydes with respect to maximal conversions can therefore readily be assessed from published hydration constants.

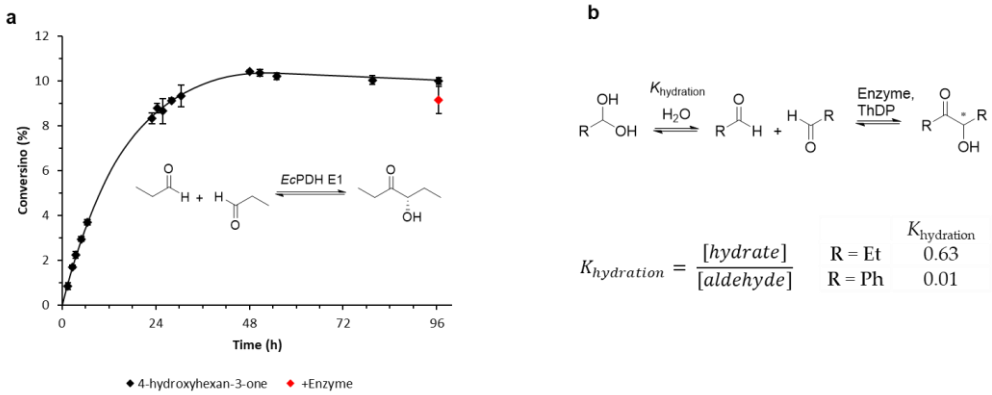


Figure 5: (a) Time course of the *Ec*PDH E1 catalysed self-reaction of propionaldehyde (2 mg/mL *Ec*PDH E1, 0.2 mM ThDP, 2 mM $MgCl_2$, 50 mM propionaldehyde, 20 mM KPi, pH 7.5). The position of the thermodynamic equilibrium limits the reaction to 10% conversion under given experimental conditions. **(b)** The hydration of aldehydes in aqueous solution introduces a competing equilibrium reaction. Electrophilic aldehydes display a large hydration constant,[56] which renders the target reaction thermodynamically less favourable.

Curiously, an enantiomeric excess of only 67% was obtained after 24 hours for the synthesis of **5b** under thermodynamic control. Conversely, decarboxylation of 2-oxobutyrate as substrate analogue afforded **5b** with 95% *ee* under kinetic control, using the same catalyst loading and reaction time. Enzyme catalysed racemization therefore seems an unlikely explanation for this observation. Since propionaldehyde functions as acceptor substrate in both cases, the disparity in stereocontrol was assumed to arise from the two different mechanisms of formation for the covalent intermediate.[5,6]

We therefore analysed the crystal structure of *holo-Ec*PDH E1 (2iea.pdb) and created models for possible configurations of the covalent intermediate *in silico* (Figure 6). The covalent bond between the ThDP cofactor and the donor substrate has double bond character through resonance, which impairs its rotation and favours a more planar configuration (see Figure 1 for the ketol intermediate in TK).

The requirement to accommodate a bulky lipoate moiety as acceptor prevents the recognition of small donor substrates *via* sterical constraints, and *Ec*PDH E1 presumably evolved to guide the approach of pyruvate towards the cofactor *via* ionic interactions. Two histidine residues (H106 and H142) are located on the left side and allow for interactions with negatively charged ketoacids. Such a prearrangement would preferentially give rise to the (*Z*)-conformation in the covalent intermediate (Figure 6a).

Our model suggests that the smaller and uncharged propionaldehyde could approach the cofactor on different trajectories, that would give rise to both (*Z*)- and (*E*)-conformers (Figure 6b). The substantial geometrical difference between the (*E*)- and (*Z*)-configurations would plausibly result in different orientations of the acceptor substrate towards the carbanion, leading to the formation of both (*R*)- and (*S*)-configured products. The synthetic performance of thermodynamically controlled reactions catalysed by ThDP dependent enzymes may therefore not only differ with respect to maximum yields, but also the obtained enantiomeric purities.

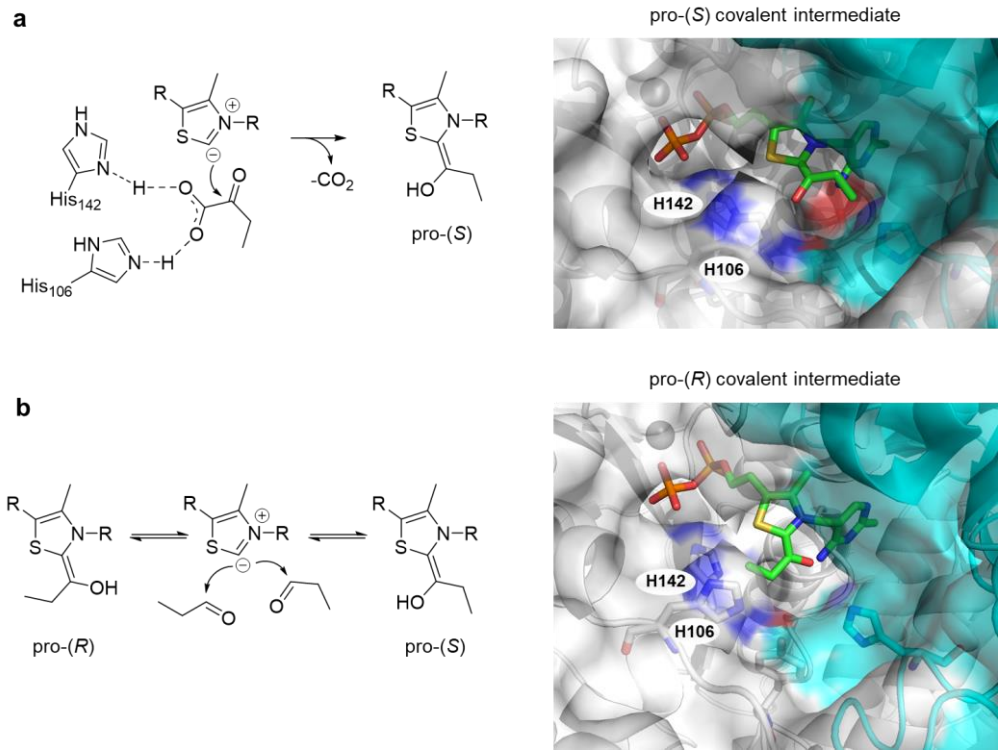


Figure 6: (a) The presence of two histidine residues (H106 and H142, surface polarity shown in blue) presumably guides the approach of negatively charged ketoacid substrates into the active side. This prearrangement could lead to the preferential formation of the (*Z*)-configured intermediate upon decarboxylation. **(b)** Modelling suggests that the smaller substrate propionaldehyde can approach the cofactor on different trajectories, which may lead to the formation of both (*E*)- and (*Z*)-configured intermediates. This could influence the orientation of the acceptor substrate and give rise to both (*R*) and (*S*)-configured acyloin products. Models and figures were created with PyMol from 2iea.pdb.

Determination of kinetic parameters

In order to compare the catalytic performance of WT *Ec*PDH E1 with other (engineered) ThDP-dependent enzymes, kinetic parameters were determined for different donor and acceptor substrates by HPLC (Table 2). Wild-type *Ec*PDH E1 showed a 250-fold higher catalytic efficiency for the conversion of pyruvate than the best engineered variant of TK_{GST} after two rounds of iterative site-saturation mutagenesis.[40] Notably, this was also 16-fold higher than the catalytic efficiency of WT *Ec*DXS, which similarly uses pyruvate as its natural donor substrate.[40] For 2-oxobutyrate, WT *Ec*PDH E1 showed a 180-fold higher catalytic efficiency than the TK_{GST} double variant H102L/H474S. Conversely, E1 proved to be a poor catalyst for the conversion of hydroxypyruvate, where it was outperformed by (engineered) variants of transketolase by three orders of magnitude.

Regarding the acceptor substrates, transketolases exhibit features for the recognition of phosphorylation and the configuration of hydroxyl groups in their natural substrates,[43,53,54] which previously needed to be removed by mutagenesis for the efficient conversion of non-natural substrates.[5,11,12,57] Here, WT *Ec*PDH E1 displayed a specific activity towards both α-hydroxyaldehydes and aliphatic aldehydes that was comparable to the engineered transketolase variants.

Notably, different variants were evolved at varying positions for the conversion of either aliphatic donor or acceptor substrates, respectively, which may not be mutually compatible. In summary, transketolases naturally evolved towards highly polar, phosphorylated carbohydrates. This makes their wild-type scaffolds inept for the conversion of aliphatic ketoacids. While moderate improvements can be achieved by rational mutagenesis, these variants are still considerably outperformed by wild-type enzymes such as *Ec*DXS[40] and *Ec*PDH E1 (this work).

Table 2: Comparison of apparent kinetic parameters for different enzymes and substrates.

Entry	Enzyme	Variant	U/mg	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}mM^{-1}$)	Ref
	TK_{GST}	H102L/H474S	n.a.[g]	16.6[a]	0.17[a]	0.01[a]	[40]
1a	*Ec*DXS	WT	n.a.[g]	3.3[a]	0.50[a]	0.15[a]	[40]
	*Ec*E1	WT	2.95[a]	4.2[a]	10.26[a]	2.45[a]	[f]
	TK_{GST}	WT	n.a.[g]	2.3[a]	12.7[a]	5.50[a]	[40]
HPA	*Sc*TK	R528Q/S527T	0.44[a]	53[a]	n.a.[g]	n.a.[g]	[12]
	*Ec*E1	WT	0.24[a]	100[a]	0.83[a]	0.008[a]	[f]
	TK_{GST}	H102L/H474S	0.006[e]	3.3[a]	0.16[a]	0.048[a]	[40,41]
1b	*Ec*E1	WT	2.72[a]	1.09[a]	9.46[a]	8.68[a]	[f]
	*Ec*E1	WT	2.33[e]				[f]
	*Ec*TK	H461S	3.14[d]	n.a.[g]	n.a.[g]	n.a.[g]	[11]
2	TK_{GST}	L191I	4.07[d]	n.a.[g]	n.a.[g]	n.a.[g]	[15]
	*Ec*E1	WT	3.09[b]	25.2[b]	10.7[b]	0.42[b]	[f]
	TK_{GST}	L191I/D470L	1.48[d]	n.a.[g]	n.a.[g]	n.a.[g]	[15]
4	TK_{GST}	H102L/H474S/ F435I	0.007[c]	n.a.[g]	n.a.[g]	n.a.[g]	[41]
	*Ec*E1	WT	2.33[c]	42.4[c]	8.10[c]	0.19[c]	[f]
8	*Sc*TK	D477E	n.a.[g]	66[d]	0.6[d]	0.009[d]	[5]
	*Ec*E1	WT	0.78[b]	69.4[b]	2.70[b]	0.039[b]	[f]

In combination with: [a] glycolaldehyde, [b] pyruvate, [c] 2-oxobutyrate, [d] hydroxypyruvate, [e] propionaldehyde. [f] this work. [g] n.a. = data not available.

Mutagenesis

While TK_{GST} variants showed some promiscuity towards the sterically more challenging donor substrate 3-methyl-2-oxobutyrate (0.004 - 0.012 $s^{-1}mM^{-1}$),[40] no conversion was observed with WT *Ec*PDH E1. Due to the absence of a crystal structure for TK_{GST} a homology model was created in YASARA. This model was then compared to the crystal structure of WT *Ec*PDH E1 (2iea.pdb, Figure 7).

Both active sites are largely comprised of the same conserved residues, but their overall structural dimensions differ considerably. Most notably, distal tertiary and quaternary structural features place the loop bearing residue H68 in TK (corresponding to H106 in E1) deeper into the active site of TK. This is not the case in *Ec*PDH E1, which allows its active site to be overall more spacious (Figure 1, Figure 7). Notably, residue H102 in TK_{GST} is replaced by Y177 in *Ec*PDH E1 as the most prominent difference in active site residues. The bulkiness of Y177 could also provide a plausible explanation for why the branched substrate 3-methyl-2-oxobutyrate was not converted by WT E1. Residue H68 is essentially conserved in TK_{GST} due to its role in cofactor binding and activation; however, mutations at H474 (S,N) and H102 (G,L,T) enabled the conversion of aliphatic ketoacids by TK_{GST}.[40]

Most of these variants are capable of forming hydrogen bonds *via* their amino-acid side chains in order to retain essential hydrogen bond interactions within the active site. Following this train of thought, we therefore chose to introduce serine residues as the smallest hydrogen bond donors at the corresponding positions (Y177S, H640S and their combination to Y177S/H640S) in *Ec*PDH E1. While all three variants retained their activity towards pyruvate, H640S no longer showed activity towards 2-oxobutyrate, and none of the three variants displayed any observable promiscuity towards either branched 3-methyl-2-oxobutyrate or linear 2-oxovalerate. While *Ec*PDH E1 constitutes a good scaffold for the conversion of 2-oxobutyrate, extensive site-saturation mutagenesis is required to further broaden its substrate scope. The expansion of the donor substrate scope of ThDP-dependent enzymes therefore remains a challenging field.

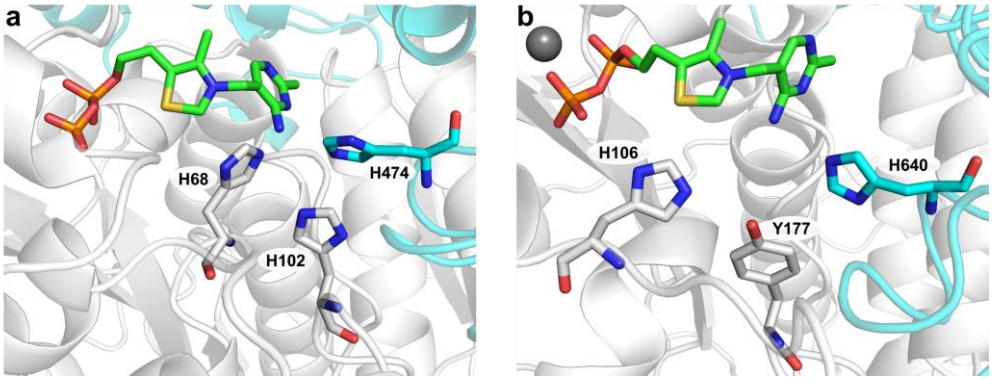


Figure 7: (a) Active site view of the homology model for TK_{GST}. Residue H68 is essential for the activation of ThDP, whereas residues H102 and H474 were successfully mutated towards the conversion of aliphatic ketoacids.[40] **(b)** Active site view of *Ec*PDH E1 (2iea.pdb). The position of H102 is occupied by Y177 as the most prominent difference between the active sites of TK_{GST} and *Ec*PDH E1. Residues Y177 and H640 were therefore similarly targeted for mutagenesis, based on the most successful variants that were previously identified in TK_{GST}. The homology model was created with YASARA, figures were created with PyMol.

3.3 Conclusions

In summary, our results show that WT *Ec*PDH E1 is a promising biocatalyst for the chiral synthesis of diverse acyloins under kinetically controlled conditions. The enzyme can be prepared in a high volumetric yield of 350 mg pure protein per litre of expression medium under non-optimised conditions, and showed excellent storage stability. *Ec*PDH E1 displayed a 180- to 250-fold higher catalytic efficiency than engineered transketolase variants for the conversion of pyruvate and 2-oxobutyrate, respectively. Additionally, *Ec*PDH E1 consistently showed a good stereoselectivity towards both aliphatic and hydroxyaldehyde substrates. The use of lipoate as its natural acceptor substrate requires an unusually broad active site

cleft, that permits the efficient conversion of both linear and branched aliphatic aldehydes, next to polar α-hydroxyaldehydes. Notably, the absence of a stereopreference with respect to the configuration of α-hydroxyaldehydes facilitates the conversion of both enantiomers, leading to a broader product scope. Taken together, these features warrant the practical application of *Ec*PDH E1 for carboligation reactions with pyruvate and 2-oxobutyrate. While thermodynamically controlled self-reactions are 100% atom efficient, large hydration constants render this approach economically and environmentally unviable for most non-aromatic aldehyde substrates. In addition, a lower level of stereocontrol was observed for the self-reaction of propionaldehyde with *Ec*PDH E1. Thermodynamically controlled reactions with ThDP dependent enzymes therefore not only deviate with respect to maximum yields from their kinetically controlled analogues, but may also differ in the enantiomeric purity of the product.

3.4 Methods summary

pH dependent activity (n=3):[52] 2,6-dichloroindophenol (DCPIP) was dried under reduced pressure in a desiccator over night before preparation of a stock solution (10 mM, in 20 mM KH_2PO_4, pH = 7.0). The isosbestic point (517 nm) was determined by measuring absorbance spectra at three different pH values (3.8, 6.3 and 8.3) and a calibration curve was prepared over a range from 0-100 µM (517 nm: ε = 3.74 $mM^{-1}cm^{-1}$, R^2 = 0.999). Reaction mixtures containing 0.2 mM ThDP, 2 mM $MgCl_2$, 2 mM sodium pyruvate and 100 µM DCPIP were incubated at 25°C. The reaction was initiated by the addition of 100 µg of holo-*Ec*PDH E1 and the decrease in absorbance was followed at the isosbestic point (517 nm, 500 rpm) in triplicate in polystyrene cuvettes on a Cary60 UV-Vis spectrometer (Agilent Technologies) equipped with a TC1 stirring unit (Quantum Northwest).

Temperature dependent activity (n=3):[52] temperature dependent specific activities were determined using an adapted version of the above DCPIP assay. The reaction mixture was incubated at the target temperature for 5 minutes before the addition of holoenzyme. The change in absorbance was recorded at the absorbance maximum (605 nm, ε = 9.01 $mM^{-1}cm^{-1}$, R^2 = 0.999) for solutions of DPCIP at pH 7.5 for an enhanced sensitivity. Aliquots of the holoenzyme were incubated at the respective temperatures (30°C, 120 min.; 40°C, 145 min.; 50°C, 150 min., 500 rpm) and their respective initial rates were determined in the same manner.

General RP-HPLC method: Samples were quenched by 1:1 dilution with 0.2% (v/v) aqueous trifluoroacetic acid, and precipitated protein was removed by centrifugation (13'000 rfc, 2 min). Samples were analysed by RP-HPLC using an ICSep ICE Coregel-87H3 column (0.4x25 cm, Transgenomic) with 0.1% v/v trifluoroacetic acid as mobile phase, 0.8 mL/min, 60°C, detection at 210 nm) on a Shimadzu LC-20AD system. Concentrations were determined using external standards.

Chiral GC methods: chiral separation of α-hydroxyketone enantiomers was achieved by chiral phase GC using previously developed methods:[58] 1 µL of sample (split 1/150) was analysed using a chiral CP-Chirasil-DEX CB column (Agilent, 25 m x 0.25 mM x 0.25 µm) on a Shimadzu 2010 Plus GC instrument equipped with an AOC-20i autosampler. Helium was used as carrier gas with a linear flow of 30 cm/sec. Injector temperature: 250°C, FID-detector temperature: 275°C.

Michaelis-Menten analysis (n=2): Reaction mixtures were prepared in potassium phosphate buffer (20 mM, pH = 7.5) containing 0.2 mM ThDP, 2 mM $MgCl_2$ and 0.1 mg/mL of *Ec*PDH E1. The substrate and enzyme solutions were brought to the target temperature (37°C, 5 min) before they were mixed to initialise the reaction in final volume of 1 mL. Reaction times were chosen to remain below 20% conversion for credible initial rate

conditions. Samples were analysed by RP-HPLC. Curve fitting was performed with the programme Igor using the Michaelis-Menten equation.

Time course for the conversion of racemic DL-glyceraldehyde with pyruvate: *Ec*PDH E1 (0.2 mg) was incubated with ThDP (0.2 mM) and $MgCl_2$ (2 mM) in potassium phosphate buffer (2 mL final volume, 20 mM KPi, pH 7.5) for 5 minutes, after which sodium pyruvate (50 mM) and DL-glyceraldehyde (50 mM) were added. Progress of the reaction was monitored by HPLC.

Homology modelling: a homology model was built for TK_{GST} with YASARA version 19.9.17[59] using its protein sequence[40] and the following parameters as input: PSI-BLAST iterations: 3; templates: 5; E-value 0.1.

References

[1] Pohl, M.; Wechsler, C.; Müller, M. Acyloin, Benzoin, and Related Reactions. In *Science of Synthesis: Biocatalysis in Organic Synthesis 2,* Faber, K., Fessner, W.-D., Turner, N.J., Eds.; Georg Thieme Verlag: Stuttgart, 2014; pp. 93-127, ISBN 9783131741714.

[2] Kluger, R.; Tittmann, K. Thiamin diphosphate catalysis: enzymic and nonenzymic covalent intermediates. *Chem. Rev.* **2008**, *108*, 1797-1833.

[3] Tittmann, K.; Golbik, R.; Uhlemann, K.; Khailova, L.; Schneider, G.; Patel, M.; Jordan, F.; Chipman, D.M.; Duggleby, R.G.; Hübner, G. NMR analysis of covalent intermediates in thiamin diphosphate enzymes. *Biochemistry* **2003**, *42*, 7885-7891.

[4] Breslow, R. On the Mechanism of Thiamine Action. IV.1 Evidence from Studies on Model Systems. *J. Am. Chem. Soc.* **1958**, *80*, 3719-3726, doi:10.1021/ja01547a064.

[5] Marsden, S.R.; Gjonaj, L.; Eustace, S.J.; Hanefeld, U. Separating Thermodynamics from Kinetics—A New Understanding of the Transketolase Reaction. *ChemCatChem* **2017**, *9*, 1808-1814.

[6] Marsden, S.R.; Mestrom, L.; McMillan, D.G.; Hanefeld, U. Thermodynamically and Kinetically Controlled Reactions in Biocatalysis–from Concepts to Perspectives. *ChemCatChem* **2020**, *12*, 426-437.

[7] Wohlgemuth, R. C2-Ketol elongation by transketolase-catalysed asymmetric synthesis. *J. Mol. Catal. B: Enzym.* **2009**, *61*, 23-29.

[8] Cázares, A.; Galman, J.L.; Crago, L.G.; Smith, M.E.; Strafford, J.; Ríos-Solís, L.; Lye, G.J.; Dalby, P.A.; Hailes, H.C. Non-α-hydroxylated aldehydes with evolved transketolase enzymes. *Org. Biomol. Chem.* **2010**, *8*, 1301-1309.

[9] Yi, D.; Saravanan, T.; Devamani, T.; Charmantray, F.; Hecquet, L.; Fessner, W.-D. A thermostable transketolase evolved for aliphatic aldehyde acceptors. *Chem. Commun.* **2015**, *51*, 480-483.

[10] Saravanan, T.; Reif, M.-L.; Yi, D.; Lorillière, M.; Charmantray, F.; Hecquet, L.; Fessner, W.-D. Engineering a thermostable transketolase for arylated substrates. *Green Chem.* **2017**, *19*, 481-489.

[11] Hibbert, E.G.; Senussi, T.; Costelloe, S.J.; Lei, W.; Smith, M.E.; Ward, J.M.; Hailes, H.C.; Dalby, P.A. Directed evolution of transketolase activity on non-phosphorylated substrates. *J. Biotechnol.* **2007**, *131*, 425-432.

[12] Ranoux, A.; Karmee, S.K.; Jin, J.; Bhaduri, A.; Caiazzo, A.; Arends, I.W.; Hanefeld, U. Enhancement of the substrate scope of transketolase. *ChemBioChem* **2012**, *13*, 1921-1931.

[13] Smith, M.E.; Hibbert, E.G.; Jones, A.B.; Dalby, P.A.; Hailes, H.C. Enhancing and reversing the stereoselectivity of Escherichia coli transketolase via single-point mutations. *Adv. Synth. Catal.* **2008**, *350*, 2631-2638.

[14] Westphal, R.; Vogel, C.; Schmitz, C.; Pleiss, J.; Müller, M.; Pohl, M.; Rother, D. A Tailor-Made Chimeric Thiamine Diphosphate Dependent Enzyme for the Direct Asymmetric Synthesis of (S)-Benzoins. *Angew. Chem. Int. Ed.* **2014**, *53*, 9376-9379.

[15] Zhou, C.; Saravanan, T.; Lorillière, M.; Wei, D.; Charmantray, F.; Hecquet, L.; Fessner, W.D.; Yi, D. Second-Generation Engineering of a Thermostable Transketolase (TKGst) for Aliphatic Aldehyde Acceptors with Either Improved or Reversed Stereoselectivity. *ChemBioChem* **2017**, *18*, 455-459.

[16] Pohl, M.; Sprenger, G.A.; Müller, M. A new perspective on thiamine catalysis. *Curr. Opin. Biotechnol.* **2004**, *15*, 335-342.

[17] Hailes, H.C.; Rother, D.; Müller, M.; Westphal, R.; Ward, J.M.; Pleiss, J.; Vogel, C.; Pohl, M. Engineering stereoselectivity of ThDP-dependent enzymes. *FEBS J.* **2013**, *280*, 6374-6394.

[18] Giovannini, P.P.; Bortolini, O.; Massi, A. Thiamine-Diphosphate-Dependent Enzymes as Catalytic Tools for the Asymmetric Benzoin-Type Reaction. *Eur. J. Org. Chem.* **2016**, *2016*, 4441-4459.

[19] Wikner, C.; Nilsson, U.; Meshalkina, L.; Udekwu, C.; Lindqvist, Y.; Schneider, G. Identification of catalytically important residues in yeast transketolase. *Biochemistry* **1997**, *36*, 15643-15649.

[20] Wikner, C.; Meshalkina, L.; Nilsson, U.; Nikkola, M.; Lindqvist, Y.; Sundström, M.; Schneider, G. Analysis of an invariant cofactor-protein interaction in thiamin diphosphate-dependent enzymes by site-directed mutagenesis. Glutamic acid 418 in transketolase is essential for catalysis. *Biol. Chem.* **1994**, *269*, 32144-32150.

[21] Kern, D.; Kern, G.; Neef, H.; Tittmann, K.; Killenberg-Jabs, M.; Wikner, C.; Schneider, G.; Hübner, G. How thiamine diphosphate is activated in enzymes. *Science* **1997**, *275*, 67-70.

[22] Shin, W.; Pletcher, J.; Blank, G.; Sax, M. Ring stacking interactions between thiamin and planar molecules as seen in the crystal structure of a thiamin picrolonate dihydrate complex. *J. Am. Chem. Soc.* **1977**, *99*, 3491-3499.

[23] Schellenberger, A. Sixty years of thiamin diphosphate biochemistry. *BBA-Protein Struct. M.* **1998**, *1385*, 177-186.

[24] Asztalos, P.; Parthier, C.; Golbik, R.; Kleinschmidt, M.; Hübner, G.; Weiss, M.S.; Friedemann, R.; Wille, G.; Tittmann, K. Strain and near attack conformers in enzymic thiamin catalysis: X-ray crystallographic snapshots of bacterial transketolase in covalent complex with donor ketoses xylulose 5-phosphate and fructose 6-phosphate, and in noncovalent complex with acceptor aldose ribose 5-phosphate. *Biochemistry* **2007**, *46*, 12037-12052.

[25] Lüdtke, S.; Neumann, P.; Erixon, K.M.; Leeper, F.; Kluger, R.; Ficner, R.; Tittmann, K. Sub-ångström-resolution crystallography reveals physical distortions that enhance reactivity of a covalent enzymatic intermediate. *Nat. Chem.* **2013**, *5*, 762-767.

[26] Kochetov, G.; Philippov, P.; Razjivin, A.; Tikhomirova, N. Kinetics of reconstruction of holo-transketolase. *FEBS Lett.* **1975**, *53*, 211-212.

[27] Kovina, M.V.; Kochetov, G.A. Cooperativity and flexibility of active sites in homodimeric transketolase. *FEBS Lett.* **1998**, *440*, 81-84.

[28] Kochetov, G.; Sevostyanova, I. Binding of the coenzyme and formation of the transketolase active center. *IUBMB Life* **2005**, *57*, 491-497.

[29] Sundström, M.; Lindqvist, Y.; Schneider, G. Three-dimensional structure of apotransketolase flexible loops at the active site enable cofactor binding. *FEBS Lett.* **1992**, *313*, 229-231.

[30] Kochetov, G.A.; Solovjeva, O.N. Structure and functioning mechanism of transketolase. *BBA-Proteins Proteom.* **2014**, *1844*, 1608-1618.

[31] Heinrich, P.C.; Steffen, H.; Janser, P.; Wiss, O. Studies on the reconstitution of apotransketolase with thiamine pyrophosphate and analogs of the coenzyme. *Eur. J. Biochem.* **1972**, *30*, 533-541.

[32] Schröder-Tittmann, K.; Meyer, D.; Arens, J.; Wechsler, C.; Tietzel, M.; Golbik, R.; Tittmann, K. Alternating sites reactivity is a common feature of thiamin diphosphate-dependent enzymes as evidenced by isothermal titration calorimetry studies of substrate binding. *Biochemistry* **2013**, *52*, 2505-2507.

[33] Seifert, F.; Golbik, R.; Brauer, J.; Lilie, H.; Schröder-Tittmann, K.; Hinze, E.; Korotchkina, L.G.; Patel, M.S.; Tittmann, K. Direct kinetic evidence for half-of-the-sites reactivity in the E1 component of the human pyruvate dehydrogenase multienzyme complex through alternating sites cofactor activation. *Biochemistry* **2006**, *45*, 12775-12785.

[34] Wilkinson, H.C.; Dalby, P.A. Novel insights into transketolase activation by cofactor binding identifies two native species subpopulations. *Sci. Rep.* **2019**, *9*, 1-13.

[35] Kovina, M.; Selivanov, V.; Kochevova, N.; Kochetov, G. Kinetic investigation of cooperativity in coenzyme binding by transketolase active sites. *Biochemistry* **1998**, *63*, 988-995.

[36] Jordan, F.; Nemeria, N.S.; Sergienko, E. Multiple modes of active center communication in thiamin diphosphate-dependent enzymes. *Acc. Chem. Res.* **2005**, *38*, 755-763.

[37] Nemeria, N.S.; Arjunan, P.; Chandrasekhar, K.; Mossad, M.; Tittmann, K.; Furey, W.; Jordan, F. 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.* **2010**, *285*, 11197-11209.

[38] Schapfl, M.; Baier, S.; Fries, A.; Ferlaino, S.; Waltzer, S.; Müller, M.; Sprenger, G.A. Extended substrate range of thiamine diphosphate-dependent MenD enzyme by coupling of two C–C-bonding reactions. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 8359-8372.

[39] Hernández, K.; Parella, T.; Petrillo, G.; Usón, I.; Wandtke, C.M.; Joglar, J.; Bujons, J.; Clapés, P. Intramolecular Benzoin Reaction Catalysed by Benzaldehyde Lyase from Pseudomonas Fluorescens Biovar I. *Angew. Chem. Int. Ed.* **2017**, *56,* 5304-5307.

[40] Saravanan, T.; Junker, S.; Kickstein, M.; Hein, S.; Link, M.K.; Ranglack, J.; Witt, S.; Lorillière, M.; Hecquet, L.; Fessner, W.D. Donor Promiscuity of a Thermostable Transketolase by Directed Evolution: Efficient Complementation of 1-Deoxy-d-xylulose-5-phosphate Synthase Activity. *Angew. Chem. Int. Ed.* **2017**, *56*, 5358-5362.

[41] Casajus, H.; Lagarde, A.; Leremboure, M.; De Dios Miguel, T.; Nauton, L.; Thery, V.; Fessner, W. D.; Duguet, N.; Charmantray, F.; Hecquet, L. Enzymatic Synthesis of Aliphatic Acyloins Catalyzed by Thermostable Transketolase. *ChemCatChem* **2020**. doi.org/10.1002/cctc.202001160

[42] Yu, H.; Hernández López, R.I.; Steadman, D.; Méndez-Sánchez, D.; Higson, S.; Cázares-Körner, A.; Sheppard, T.D.; Ward, J.M.; Hailes, H.C.; Dalby, P.A. Engineering transketolase to accept both unnatural donor and acceptor substrates and produce α-hydroxyketones. *FEBS J.* **2019**, *287*, 1758-1776.

[43] Kruger, N.J.; von Schaewen, A. The oxidative pentose phosphate pathway: structure and organisation. *Curr. Opin. Plant Biol.* **2003**, *6*, 236-246.

[44] Vogel, C.; Pleiss, J. The modular structure of ThDP-dependent enzymes. *Proteins: Struct. Funct. Bioinf.* **2014**, *82*, 2523-2537.

[45] Widmann, M.; Radloff, R.; Pleiss, J. The Thiamine diphosphate dependent Enzyme Engineering Database: A tool for the systematic analysis of sequence and structure relations. *BMC Biochem.* **2010**, *11*, 9.

[46] Vogel, C.; Widmann, M.; Pohl, M.; Pleiss, J. A standard numbering scheme for thiamine diphosphate-dependent decarboxylases. *BMC Biochem.* **2012**, *13*, 24.

[47] Buchholz, P.C.; Vogel, C.; Reusch, W.; Pohl, M.; Rother, D.; Spieß, A.C.; Pleiss, J. BioCatNet: A Database System for the Integration of Enzyme Sequences and Biocatalytic Experiments. *ChemBioChem* **2016**, *17*, 2093-2098.

[48] Yi, J.; Nemeria, N.; McNally, A.; Jordan, F.; Machado, R.S.; Guest, J.R. Effect of substitutions in the thiamin diphosphate-magnesium fold on the activation of the pyruvate dehydrogenase complex from Escherichia coli by cofactors and substrate. *J. Biol. Chem.* **1996**, *271*, 33192-33200.

[49] Behal, R.; Buxton, D.; Robertson, J.; Olson, M. Regulation of the pyruvate dehydrogenase multienzyme complex. *Annu. Rev. Nutr.* **1993**, *13*, 497-520.

[50] Park, Y.-H.; Wei, W.; Zhou, L.; Nemeria, N.; Jordan, F. Amino-terminal residues 1−45 of the Escherichia coli pyruvate dehydrogenase complex E1 subunit interact with the E2 subunit and are required for activity of the complex but not for reductive acetylation of the E2 subunit. *Biochemistry* **2004**, *43*, 14037-14046.

[51] Bisswanger, H. Substrate specificity of the pyruvate dehydrogenase complex from Escherichia coli. *J. Biol. Chem.* **1981**, *256*, 815-822.

[52] Ke, C.; He, Y.; He, H.; Yang, X.; Li, R.; Yuan, J. A new spectrophotometric assay for measuring pyruvate dehydrogenase complex activity: a comparative evaluation. *Anal. Methods* **2014**, *6*, 6381-6388.

[53] Humphrey, A.J.; Parsons, S.F.; Smith, M.E.; Turner, N.J. Synthesis of a novel N-hydroxypyrrolidine using enzyme catalysed asymmetric carbon–carbon bond synthesis. *Tetrahedron Lett.* **2000**, *41*, 4481-4485.

[54] Effenberger, F.; Null, V.; Ziegler, T. Preparation of optically pure L-2-hydroxyaldehydes with yeast transketolase. *Tetrahedron Lett.* **1992**, *33*, 5157-5160.

[55] Bykova, I.A.; Solovjeva, O.N.; Meshalkina, L.E.; Kovina, M.V.; Kochetov, G.A. One-substrate transketolase-catalyzed reaction. *Biochem. Biophys. Res. Commun.* **2001**, *280*, 845-847.

[56] Hilal, S.; Bornander, L.; Carreira, L. Hydration equilibrium constants of aldehydes, ketones and quinazolines. *ACS Comb. Sci.* **2005**, *24*, 631-638.

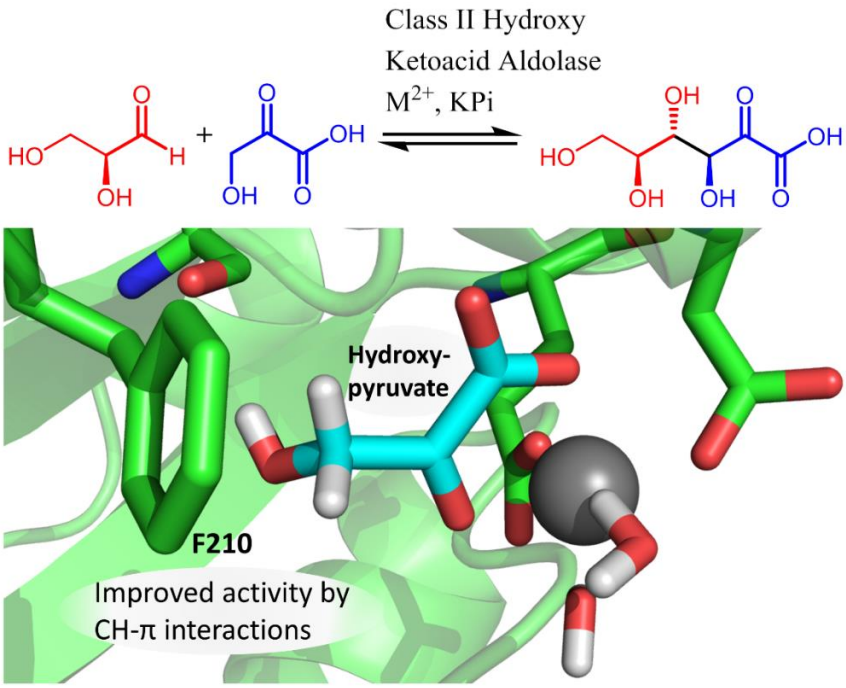
[57] Hibbert, E.G.; Senussi, T.; Smith, M.E.; Costelloe, S.J.; Ward, J.M.; Hailes, H.C.; Dalby, P.A. Directed evolution of transketolase substrate specificity towards an aliphatic aldehyde. *J. Biotechnol.* **2008**, *134*, 240-245.

[58] Médici, R.; Stammes, H.; Kwakernaak, S.; Otten, L.G.; Hanefeld, U. Assessing the stereoselectivity of Serratia marcescens CECT 977 2, 3-butanediol dehydrogenase. *Catal. Sci. Technol.* **2017**, *7*, 1831-1837.

[59] Krieger, E.; Vriend, G. YASARA View—molecular graphics for all devices—from smartphones to workstations. *Bioinformatics* **2014**, *30*, 2981-2982.

4

CH-π interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase



This chapter is based on

Stefan. R. Marsden, Luuk Mestrom, Isabel Bento, Peter-Leon Hagedoorn, Duncan G. G. McMillan and Ulf Hanefeld

Advanced Synthesis and Catalysis, **2019**, *361*, 2649-2658.

DOI:10.1002/adsc.201900205

The class II hydroxy ketoacid aldolase A5VH82 from *Sphingomonas wittichii* RW1 (*Sw*HKA) accepts hydroxypyruvate as nucleophilic donor substrate, giving access to synthetically challenging 3,4-dihydroxy-α-ketoacids. The crystal structure of holo-*Sw*HKA in complex with hydroxypyruvate revealed CH-π interactions between the C-H bonds at C-3 of hydroxypyruvate and a phenylalanine residue at position 210, which in this case occupies the position of a conserved leucine residue. Mutagenesis to tyrosine further increased the electron density of the interacting aromatic system and effected a rate enhancement by twofold. While the leucine variant efficiently catalyses the enolisation of hydroxypyruvate as the first step in the aldol reaction, the enol intermediate then becomes trapped in a disfavoured configuration that considerably hinders subsequent C-C bond formation. In *Sw*HKA, micromolar concentrations of inorganic phosphate increase the catalytic rate constant of enolisation by two orders of magnitude. This rate enhancement was now shown to be functionally conserved across the structurally distinct $(\alpha/\beta)_8$ barrel and αββα sandwich folds of two pyruvate aldolases. Characterisation of the manganese (II) cofactor by electron paramagnetic resonance excluded ionic interactions between the metal centre and phosphate. Instead, His44 was shown to be primarily responsible for the binding of phosphate in the micromolar range and the observed rate enhancement in *Sw*HKA.

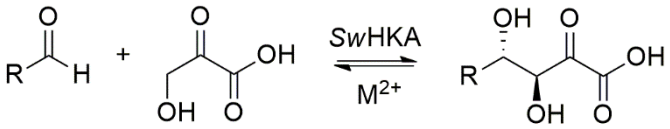
4.1 Introduction

Aldolase catalysed asymmetric C-C bond formation plays a pivotal role in the synthesis of many chiral building blocks,[1] natural compounds,[2] pharmaceuticals[3] and has been industrially applied on a multi ton scale.[4] Due to this wide-ranging utility, its relevance for synthetic chemistry is well recognized.[5]

There are two distinct types of aldolases that are classed by their respective mode of action. Class I aldolases function *via* a Schiff base mechanism involving a lysine residue,[6] while class II aldolases require a divalent metal cation as cofactor for substrate binding and activation.[7] In general, aldol reactions are fully reversible *via* the same reaction pathway. In nature, such retro-aldol reactions are found in biodegradation pathways of aromatic compounds such as phenols, biphenyls and lignin derived metabolites, which then can be metabolised.[8]

For synthetic purposes, aldolase catalysed reactions with pyruvate are well established and give straightforward access to 4-hydroxy-α-ketoacids.[1a, 9] Unfortunately, wild type pyruvate aldolases typically show a strict requirement for a single donor substrate, which considerably limits their product scope.

To address this issue, de Berardinis *et al*. conducted an extensive screening of 571 aldolases, of which only 19 enzymes showed an observable promiscuity towards hydroxypyruvate (HPA) as nucleophilic donor substrate.[10] In particular, aldolase A5VH82 from *Sphingomonas wittichii* RW1 showed good conversions, giving new access to several 3,4-dihydroxy-α-ketoacids by variation of the acceptor aldehyde. Intrigued by its unusual activity, we set out to elucidate the molecular basis behind the enzyme's widened donor substrate scope and to conduct an extensive biochemical and structural characterisation. Our findings motivated us to denote this enzyme as a hydroxy ketoacid aldolase (*Sw*HKA, Scheme 1).



Scheme 1: *Sw*HKA catalysed aldol coupling of hydroxypyruvate with an acceptor aldehyde to afford a (*3S*, *4S*) configured 3,4-dihydroxy-2-oxoacid.[10]

A first indication of phosphate being involved in the catalytic mechanism of class II pyruvate aldolases was given by the crystal structure of the 2-dehydro-3-deoxy-galactarate aldolase from *Escherichia coli* (DDG, 1dxf.pdb). Based on the apparent absence of catalytically active amino acid residues and an increased activity in phosphate buffer, the authors proposed a reaction mechanism which involved inorganic phosphate as the catalytic base.[11] This hypothesis has since been challenged by observations for the homologous enzyme HpaI (44% sequence identity with DDG) where lower activities were found in the presence of phosphate.[12] While a recent study on the class II HMG/CHA pyruvate aldolase described a 10-fold rate enhancement by inorganic phosphate, these findings were attributed to the enzyme's distinct structure (αββα sandwich fold in the HMG/CHA aldolase [13] *vs*. an $(\alpha/\beta)_8$ barrel fold in the DDG aldolase[11]; 16.8% shared sequence identity).[14]

Here we present new insights into the mode of phosphate activation by identifying the key residues that are involved in the binding of phosphate and the concomitant rate enhancement. Our results show that phosphate activation is functionally conserved within the same order of magnitude in the $(\alpha/\beta)_8$ barrel fold of *Sw*HKA and thereby not a unique property of the αββα sandwich fold. Furthermore, using the crystal structure of holo-*Sw*HKA in complex with hydroxypyruvate, we were able to demonstrate CH-π interactions between residue F210 and HPA to be responsible for its conversion. Based on this, we could then rationally devise a variant of two-fold improved turnover rate.

4.2 Results and Discussion

Biocatalyst production

The *Sw*HKA gene was codon optimised for recombinant expression in *E. coli* BL21(DE3) and a purified protein yield of up to 75 mg/L medium was obtained upon single-step metal affinity purification *via* its *N*-terminal His_6-tag. Notably, gene expression in *E. coli* BL21Star(DE3) yielded approximately 2.5 g/L of insoluble inclusion bodies with *Sw*HKA as its main constituent, while only 10 - 20 mg/L was present in its soluble form. The inclusion bodies were purified with sodium deoxycholate and subsequently re-dissolved in urea, but no functional enzyme could be recovered. Cheap and high yielding biocatalyst production from purified inclusion bodies was therefore not possible.

Overall, gene expression in *E. coli* BL21(DE3) gave higher yields of soluble protein. For crystallisation studies, the soluble enzyme was further purified by size exclusion chromatography resulting in a single peak corresponding to a hexameric homo-oligomer (27.4 kDa x 6 = 164.4 kDa), a commonly found structural state of other class II pyruvate aldolases.[11, 15]

Crystal structure of Mg^{2+}-*Sw*HKA (6r62.pdb)

The crystal structure of holo-*Sw*HKA in complex with hydroxypyruvate was determined at 1.55 Å resolution. The final structural model corresponds to a monomeric subunit and includes all 251 amino acid residues plus one histidine residue from the His_6-tag (Figure 1a).

Additionally, one Mg^{2+} ion and one hydroxypyruvate moiety were found fully occupied in the electron-density maps, together with 352 solvent molecules. The monomeric subunit shows an $(\alpha/\beta)_8$ triosephosphate isomerase (TIM) fold, similar to the structures of the homologues 4-hydroxy-2-oxo-heptane-1,7-dioate aldolase (HpaI)[15] and 2-dehydro-3-deoxygalactarate aldolase (DDG).[11]

The 8th α-helix extends out of the β-barrel and packs onto a β-sheet of a neighbouring two-fold related subunit to form a domain-swapped dimer (Figure 1b).

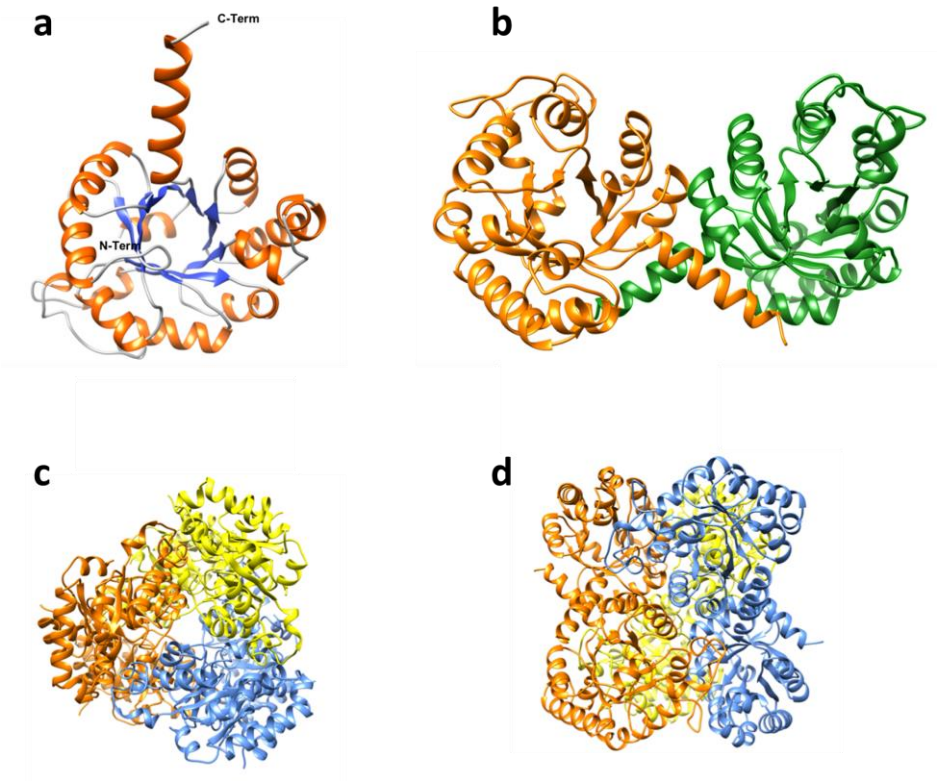


Figure 1: (a) TIM barrel fold of monomeric *Sw*HKA. **(b)** Dimerisation occurs *via* domain swapping of two α-helixes packing onto β-sheets. **(c, d)** Trimerization of domain swapped dimers affords the overall hexameric assembly of *holo*-*Sw*HKA.

In agreement with previous results from size exclusion chromatography, a final hexameric assembly was obtained by application of the 3-fold crystallographic symmetry to this domain swapped dimer. (Figure 1c,d). Comparison with the homologous enzymes HpaI (rmsd 1.54 Å for 243/251 aligned residues) and DDG (rmsd 1.61 Å for 243/251 aligned residues) indicated, that the overall fold is conserved in *Sw*HKA. The active site is located inside a cavity close to the C-terminus, and comprises of residues from two different monomeric subunits and a metal cluster (Figure 2).

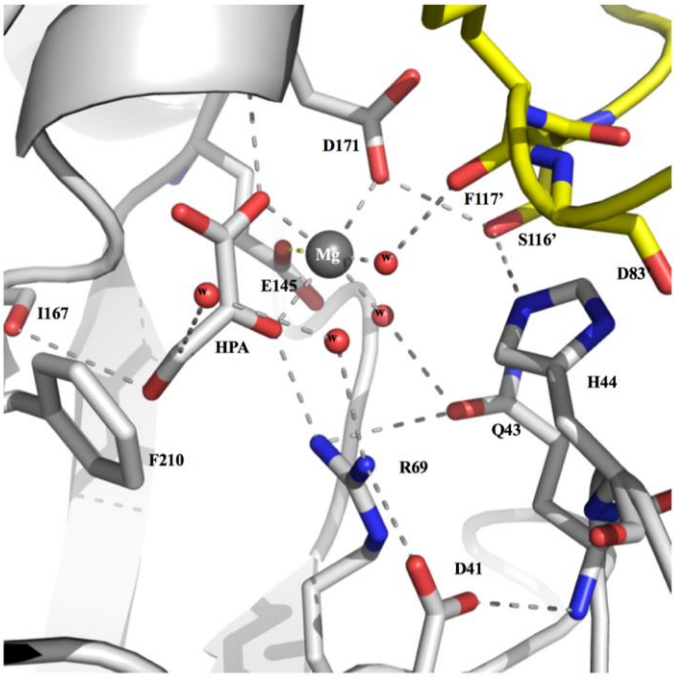


Figure 2: Representation of the relevant active site residues in Mg^{2+}-*Sw*HKA (6r62.pdb). The backbones of the two monomeric subunits that form the active site are coloured in grey and yellow. The figure was created with PyMOL.[16]

The Mg^{2+} ion is coordinated by two carboxylate groups from residues E145 and D171 in a monodentate fashion, two water molecules (W1, W2) and the bidentate binding of hydroxypyruvate via its carbonyl and the carboxylate group completes the octahedral geometry. HPA therefore coordinates the metal cofactor in a similar way as pyruvate in homologous structures.[11,15]

The carboxylate group in HPA is located within hydrogen bond distance to the backbone NH and side chain hydroxyl of T170. The carbonyl group in HPA forms a hydrogen bond with R69 and its C-3 hydroxyl group is within hydrogen bond distance of the carbonyl backbone of I167. Furthermore, a well-ordered water molecule is observed at 3 Å distance from C-3 in HPA. This water molecule

constitutes a good candidate for the catalytic deprotonation at C-3, similar to the proposed role of water in the structure of pyruvate bound HpaI (4b5u.pdb).[15] This bridging interaction is necessary due to the absence of proximal amino acid residues for a direct deprotonation. Notably, F210 replaces a conserved leucine residue (L212 and L216 for HpaI and DDG, respectively) as the most apparent difference in active site geometries. The C-3 C-H bonds in HPA are oriented towards the aromatic system of F210 and establish CH-π interactions.[17] This hydrophobic region was previously proposed to flank the aldehyde binding locus.[15] Another important and conserved residue in the active site is H44 (H45 and H50 in HpaI and DDG, respectively). H44 is located within hydrogen bond distance of D83' in the symmetry related subunit and of a water molecule that coordinates the Mg^{2+} ion. Together with the loop region that comprises of residues 115' to 120' in the symmetry related molecule, D83' is involved in a hydrogen bond network that contributes to the integrity of the active site. H44 forms hydrogen bonds with R120' and D171. Furthermore, D171 forms hydrogen bonds with F117', S116' and one of the water molecules from the Mg^{2+} coordination sphere. Although this loop shows low sequence conservation among the homologous enzymes, it maintains the active site configuration via the hydrogen bond network.

Biochemical characterisation

In class II aldolases, divalent metal ions play a central role in catalysis.[7] For *Sw*HKA, a slight preference for Mn^{2+} over Mg^{2+} and Co^{2+} was previously reported, but no dissociation constants or kinetic parameters were measured.[10]

Therefore, the metal dependent activity of the enzyme was screened, including Ni^{2+}, Ca^{2+} and Zn^{2+}, with the latter being commonly found in the dihydroxyacetone phosphate specific subgroup of class II aldolases.[18] The highest activity was observed with Mn^{2+} as cofactor, followed by Mg^{2+} with 83% of the activity of Mn^{2+} (Figure 3).

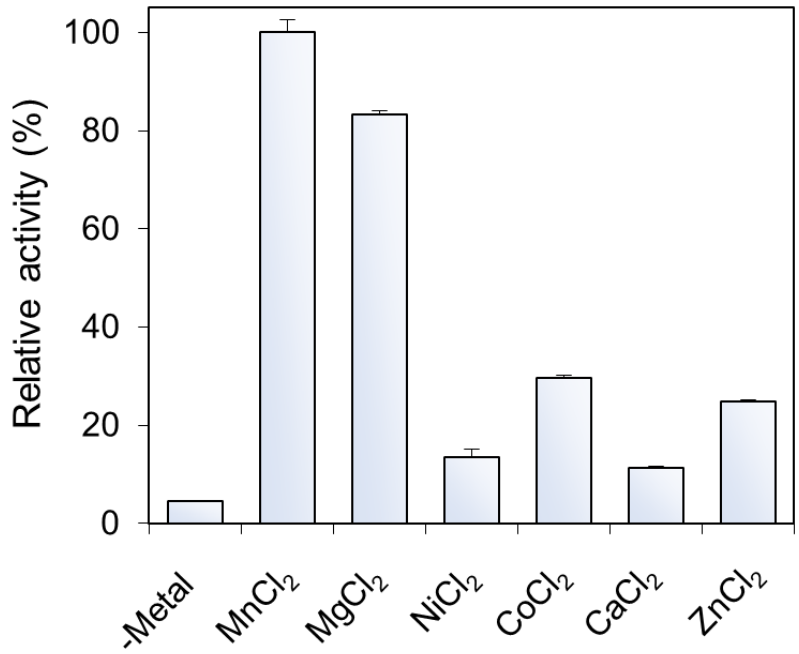


Figure 3: Metal dependent relative activities of SwHKA under metal saturated conditions (10 mM M^{2+}, 20 mM TEOA, pH 8.0). Activities were determined for the retro-aldol type decarboxylation of oxaloacetate by a NADH/LDH coupled enzyme assay and results were normalised for Mn^{2+}.

While rather similar in rate, a strong preference for Mn^{2+} over Mg^{2+} was observed with respect to their apparent binding constants, with a K_d of 3.3±0.6 µM for Mn^{2+} in contrast to 58.4±4.1 µM for Mg^{2+}. These results are of particular interest, since solvated Mg^{2+} and Mn^{2+} ions are known to be poor Lewis acid catalysts in aqueous media.[19] A previous study on the Mn^{2+} containing enzyme *Gt*HNL suggested a distorted octahedral coordination geometry to be responsible for the increased Lewis acidity, where the connecting line between the two tops of the square-based pyramid deviates from the expected angle of 180°.[20] A similar deviation was also observed in the crystal structure of holo-*Sw*HKA for the Mg^{2+} cluster (6r62.pdb).

To determine the pH optimum, the enzyme activity was measured in several buffer systems over a range from 5.5 to 8.25 and the optimum was found at pH 6.75. Notably, long-term storage of the enzyme at -20°C required alkaline conditions (pH ≥7.5), under which the enzyme retained full activity for more than one month. Incubation in potassium phosphate buffer substantially increased the reaction rate in comparison with other buffer salts. Analysis over a broad range of concentrations revealed a considerable affinity towards phosphate, with an apparent dissociation constant of $K_{d,Pi}$ = 175±36µM. This observation was of particular interest, since phosphate buffers were reported to exhibit an adverse effect on the activity of class II aldolases in several other cases,[12, 21] with only one exception.[14] Strikingly, even minor concentrations of phosphate significantly increased the reaction rate up to 13-fold under standard assay conditions, from 2.1 U/mg to v_{sat} = 27.1 U/mg following saturation kinetics with respect to phosphate.

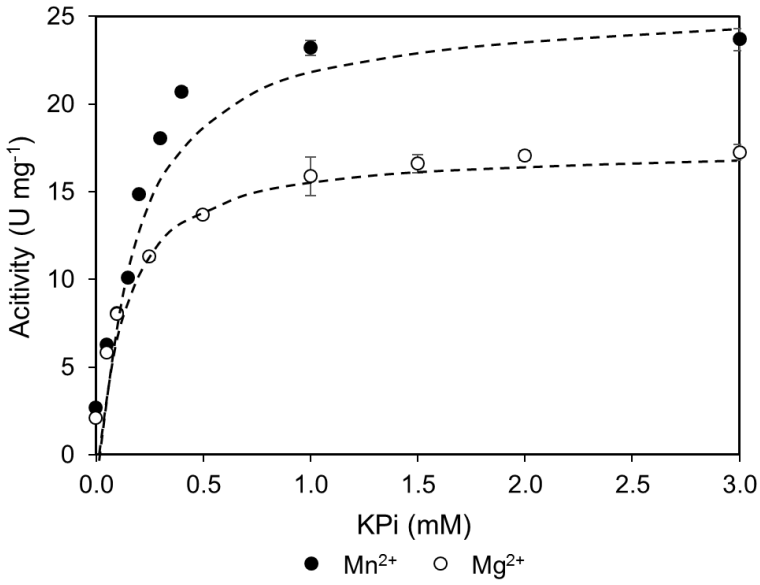


Figure 4: Catalytic rate enhancement in SwHKA by the presence of inorganic phosphate under metal saturated conditions (0.1 mM Mn^{2+} or 1 mM Mg^{2+}, 5 mM TEOA, pH 7.5). Rates were determined by the coupled enzyme assay.

The metal dissociation constants for Mn^{2+} and Mg^{2+} were unaffected by the presence/absence of phosphate, rendering a direct ionic interaction unlikely (Table 1). Henceforth, we directed our attention towards the two positively charged active site residues R69 and H44, which appeared to be likely candidates involved in phosphate binding. While H44 also plays a structural role by interacting with D83 from another dimer, mutant H44A still displayed 40% (0.78±0.01 U/mg) residual activity in the absence of phosphate.

However, the mutant's affinity towards phosphate was reduced 36-fold with $K_{d,Pi}$ = 6.22±0.53 mM and the related rate enhancement was only 3-fold (2.23±0.08 U/mg). Since both phosphate binding and activation are not completely lost by a replacement with alanine, H44 seems to be primarily, but not solely responsible for the observed activation of *Sw*HKA by inorganic phosphate.

In agreement with previous studies, a complete loss of activity was observed for the R69A mutant both in the absence and presence of saturating concentrations of phosphate. These findings demonstrate that neither H44 in combination with P_i, nor phosphate alone can act as the catalytic base and that R69 is required.

Despite a low dissociation constant, measurements of inorganic phosphate after metal affinity chromatography did not support its enzyme bound co-purification, indicating that while beneficial, phosphate is not essential for catalysis. These results are consistent with previous studies, where a structural molecule of water was identified as the catalytic base, due to the lack of proximal amino acid residues for a direct deprotonation at C-3 in the substrate.[9a, 12, 15, 22] Ultimately, elevated concentrations of phosphate (>10 mM KPi) were found to have a negative impact on the reaction rate. We propose this to be due to the typically low solubility of divalent metal phosphates, which deprives the active site from its metal cofactor.[23]

Table 1: Overview of apparent metal or phosphate dissociation constants and the corresponding saturation kinetics v_{sat} under standard assay conditions (0.5 mM oxaloacetate).

Enzyme	M^{2+}	Pi	v_{sat} (U/mg)	$K_{d,M2+}$ (µM)	$K_{d,Pi}$ (mM)
WT	Mn^{2+}	-	2.15±0.09	3.3±0.6	-
WT	Mg^{2+}	-	1.65±0.03	58.4±4.1	-
WT	Mn^{2+}	+	27.0±0.67	6.5±0.7	0.18±0.04
WT	Mg^{2+}	+	17.21±0.29	158±12	0.13±0.02
H44A	Mn^{2+}	-	0.78±0.01	n.d.	-
H44A	Mn^{2+}	+	2.23±0.08	n.d.	6.22±0.53

Notably, a higher rate of reaction can still be observed for states where the apoenzyme predominates, due to the rate enhancing effect of phosphate. This issue was addressed experimentally by using the minimal concentrations of metal and phosphate that are required for establishing saturating conditions. These findings plausibly clarify the previous misconception of the effect of phosphate in literature, where it typically was added in a large excess of 50 - 1000 molar equivalents with respect to the metal.[12, 21, 22d]

Finally, we would like to mention that amine buffers are known organocatalysts for the transketolase-like conversion of hydroxypyruvate.[24] Using decarboxylation as driving force, this reaction then affords a racemic dihydroxyketone in place of the desired aldol product.[25] This considerably limits the range of buffer systems which would allow for selective aldol reactions with hydroxypyruvate at neutral pH. In theory, structural analogues of phosphate with

matching pKa values should also be able to act as the catalytic base. Indeed, AsO_4^{3-} and VO_4^{3-} were previously shown to similarly activate the HMG/CHA aldolase, while SO_4^{2-} and MoO_4^{2-} did not.[14]

EPR characterisation of the metal centre

To further elucidate the interplay between the metal cofactor, inorganic phosphate and the substrates, the metal centre was analysed by EPR (Figure 5). Mn^{2+} is paramagnetic, characterised by a high spin S = 5/2 and the hyperfine coupling of the nuclear spin I = 5/2 of ^{55}Mn (100% natural abundance) and was therefore chosen as the metal cofactor for EPR experiments. The Mn^{2+} EPR is characteristic, as it is a unique case where the zero-field splitting (which is independent of the magnetic field) is smaller or equal to the Zeeman interaction (dependent on the magnetic field). This is contrary to most other paramagnetic metal ions, for which the zero field splitting is much larger, e.g. high spin Fe^{3+} (S = 5/2). The spectrum is very broad because of five allowed transitions with each $\Delta m_s = 1$, and each of these transitions is split into six lines due to the ^{55}Mn hyperfine coupling. Normally, Mn^{2+} is dominated by the $m_s = -1/2$ to $m_s = +1/2$ transition, which results in a broad signal around g = 2 with the characteristic splitting into six lines. Additional lines result from semi-forbidden transitions that have a lower intensity. A more extensive description of Mn^{2+} EPR of proteins has been published previously.[20]

The EPR spectrum of Mn^{2+} bound aldolase is dominated by the characteristic six line pattern around g = 2, with a hyperfine coupling constant of circa 95 Gauss. This hyperfine coupling constant is consistent with an octahedral Mn^{2+} complex with mixed oxygen ligands.[26] Association of potentially bidentate binding aldehyde acceptors (glycolaldehyde, methylglyoxal) did not significantly alter the EPR spectrum of the holoenzyme, while the donor substrates pyruvate and hydroxypyruvate considerably changed the EPR signal (Figure 5, top).

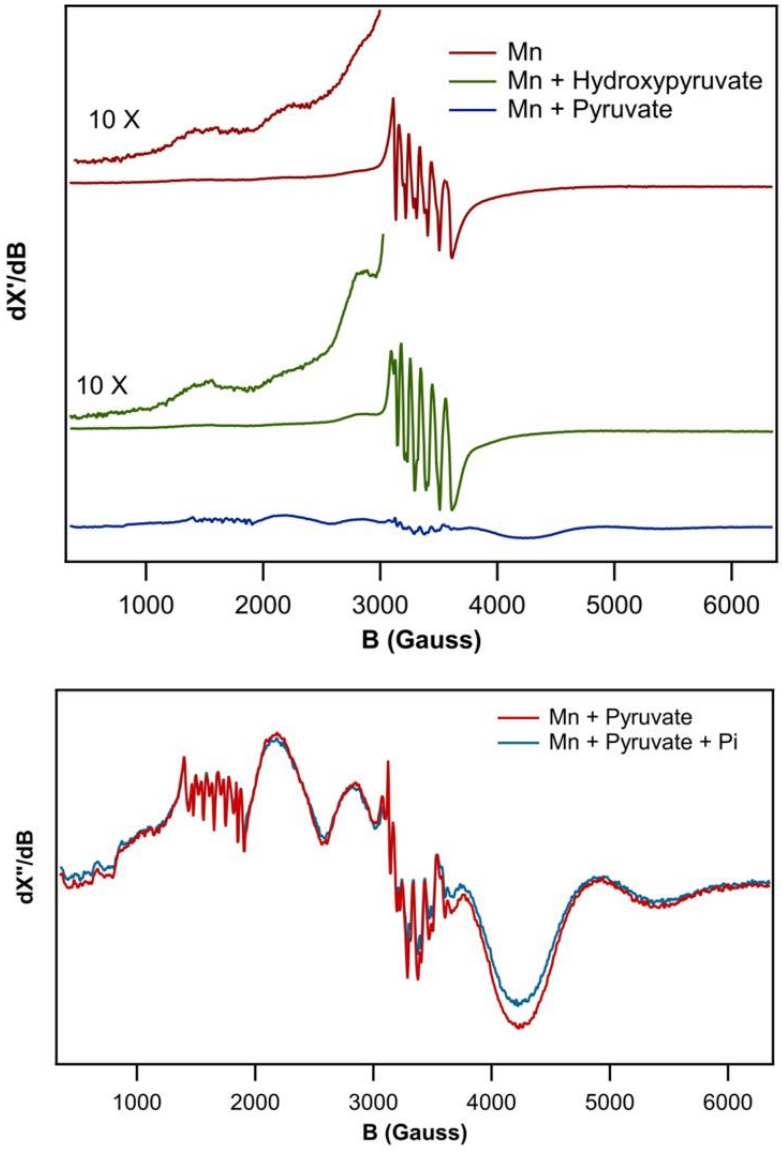


Figure 5: EPR spectroscopy of Mn(II) bound *Sw*HKA, showing the effect of donor substrates on the spectrum of Mn^{2+}-SwHKA (top) and no effect by the addition of KPi (bottom). EPR conditions: microwave frequency, 9.402 GHz; microwave power, 0.2 mW; modulation frequency, 100 kHz; modulation amplitude, 10 Gauss; temperature, 37K.

The ketoacid functional group is therefore required for coordination and activation of the donor substrate. Pyruvate dramatically altered the spectrum, as the higher order spin transitions with $\Delta m_s = 1$ became dominant. The EPR spectrum of pyruvate bound Mn-*Sw*HKA exhibits features at g = 1.13, 1.44 2.16, 2.81 and 4.3 that represent higher order spin transitions. The feature around g = 4.3 exhibits the characteristic 6-line pattern with a hyperfine coupling constant of circa 95 Gauss. This is indicative of a large change in the electronic structure of the Mn^{2+} ion, which would result from a change in the coordination number and geometry.

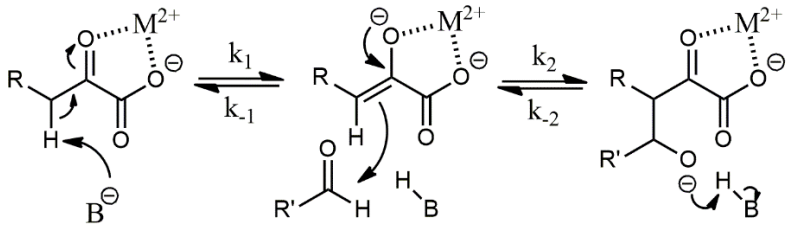
The change after addition of hydroxypyruvate is subtler, but significant and in agreement with the bidentate coordination of hydroxypyruvate in the crystal structure of Mg^{2+}-*Sw*HKA. Most importantly, addition of phosphate to the holoenzyme both in the presence and absence of substrate did not alter the EPR spectra (Figure 5, bottom), which confirmed our previous notion that there is no direct interaction between the metal and phosphate in the active site.

A similar change in the zero-field splitting has previously been reported for the Mn(II) bound extradiol cleaving catechol dioxygenase upon the addition of the substrate 3,4-dihydroxyphenylacetate and for Mn(II) bound pyruvate kinase upon the addition of pyruvate or phosphoenolpyruvate.[27] Spectra similar to the pyruvate bound Mn-*Sw*HKA have also been reported for oxalate decarboxylase, oxalate oxidase, hydroxynitrile lyase *Gt*HNL and phosphoglucosemutase.[20, 28]

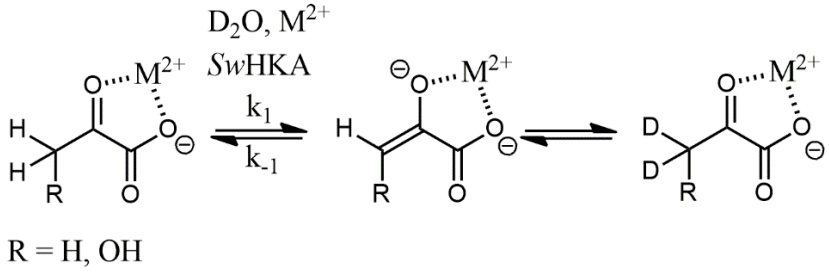
Kinetic analysis of phosphate activation

On a mechanistic level, the aldol reaction between (hydroxy-)pyruvate and an aldehyde acceptor is fully reversible *via* the same reaction pathway and comprises of three distinct reaction steps, starting from the bound substrate (Michaelis complex). These steps are characterised by their corresponding rate constants ($k_1 – k_3$ and $k_{-1} – k_{-3}$).

The reaction is initialised by deprotonation at C-3 in the donor substrate (k_1) to afford an enolate intermediate. Nucleophilic attack on the aldehyde then creates a new C-C bond (k_2) and final protonation of the alkoxide closes the catalytic cycle (k_3, for simplicity not shown).

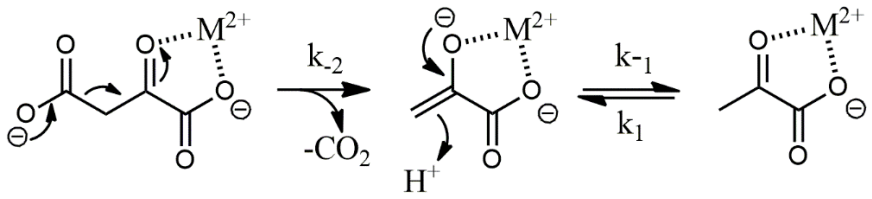


Scheme 2: Discrete reaction steps during the aldol reaction of (hydroxy-) pyruvate and an acceptor aldehyde. In the absence of aldehydes in deuterated buffers, the enzyme catalyses the H-D exchange in the donor substrate. This allows to study the impact of phosphate on the kinetic rate constants of (de-) protonation (k_1 and k_{-1}).



Scheme 3: *Sw*HKA catalysed H-D exchange in (hydroxy-) pyruvate in deuterated buffers using the non-paramagnetic Mg^{2+} as metal cofactor. This allows to specifically study enolisation as the first step in the aldol reaction.

The aldolase catalysed decarboxylation of oxaloacetate then represents an extended model system, which also takes the C-C bond forming/breaking step (k_2 and k_{-2}) into account. In this case, irreversible C-C bond breaking (decarboxylation, K_{eq} = 1.36 × 10^6 at pH 7.5) effects the direct formation of the active enolate species. Subsequent protonation of the enolate intermediate (k_{-1}) then affords the final product pyruvate.



Scheme 4: *Sw*HKA catalysed retro-aldol type decarboxylation of oxaloacetate. C-C bond breaking (k_{-2}) directly affords the enol intermediate and the subsequent protonation step corresponds to rate constant k_{-1}.

Remarkably, the presence of phosphate increases the initial rate of H-D exchange in pyruvate by 120-fold. In contrast, the K_M for oxaloacetate was unaffected by the presence of phosphate, while the catalytic rate constant k_{cat} only increased 10-fold. Phosphate therefore does not improve the binding of the substrate, but exclusively increases its rate of decarboxylation.

In summary, these findings indicate that when oxaloacetate is used as substrate, (de-)protonation is rate-limiting in the absence of phosphate, while C-C bond formation/breaking becomes rate limiting in its presence. For acceptor substrates, where electronic effects or steric constraints do not already render C-C bond formation rate limiting to begin with, a theoretical rate enhancement of up to 120-fold could be achieved by increasing the rate of enolisation when phosphate saturated conditions are used.

CH-π interactions

Since *Sw*HKA was previously shown to be particularly good at utilising HPA in aldol reactions, its active site geometry was compared to a previously published model of active site conservation among the 19 class II pyruvate aldolases that were found to be active towards HPA.[10] While highly similar, residue F210 in *Sw*HKA occupies the position of a generally conserved leucine residue as the sole apparent deviation from the model. A similarly positioned phenylalanine residue was previously reported for the class I KDPG aldolase, where the authors described its function to cap the pyruvate methyl group.[29]

In carbohydrate utilising enzymes, CH-π interactions between aromatic residues and electropositive saccharide C-H bonds have been extensively investigated and are considered one of the key determinants for carbohydrate recognition in proteins.[17, 30] Briefly, electron rich aromatic systems donate electron density to electropositive hydrogen atoms, thereby establishing binding interactions that compensate for the mismatch in polarity. In analogy to carbohydrate utilising enzymes, we therefore hypothesised that CH-π interactions between F210 and the C-H bonds at C-3 in HPA are responsible for its efficient conversion by *Sw*HKA. To provide evidence for our hypothesis, we created the variant F210L as a representative of the conserved active site geometry of pyruvate aldolases[10] and two variants F210Y and F210W which possess higher electron densities in their aromatic system than phenylalanine.[17]

To our surprise, all variants were comparably able to catalyse the formation of the enol intermediate from hydroxypyruvate during H-D exchange experiments (Figure 5), where variant F210L showed 50% and F210Y 175% relative activity with respect to the WT. The F210W mutant was only slightly better than F210L and showed 60% relative activity, which could either be due to its steric constraints, or binding interactions which are too strong for efficient turnover.

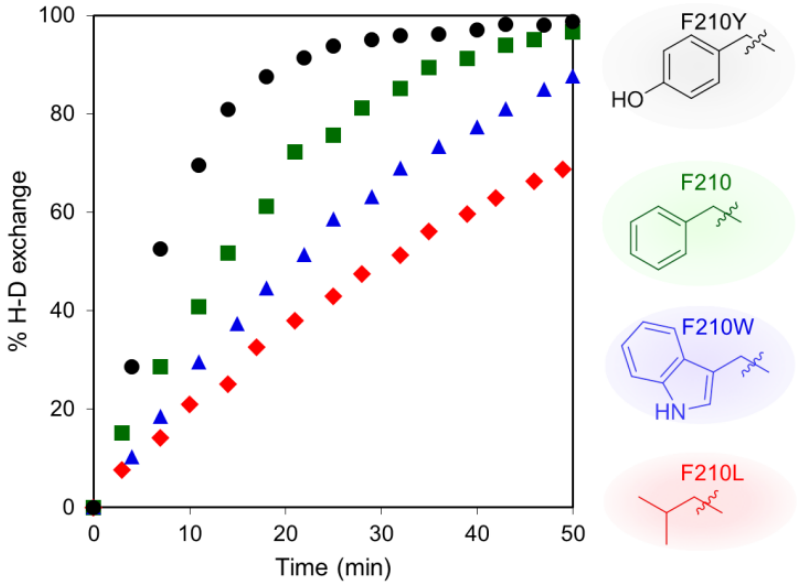


Figure 6: *Sw*HKA catalysed H-D exchange in hydroxypyruvate. F210X mutants in deuterated buffer (5 mM KPi, pD = 7.4). Control experiments in the absence of enzyme sowed <1% of background exchange over the investigated time frame.

Nevertheless, substantial differences were observed when the previously reported conversion of hydroxypyruvate with DL-glyceraldehyde[10] was followed over time (Figure 7). Equilibrium conditions were reached within 30 minutes and 1.5 hours for the F210Y mutant and the WT respectively, whereas the F210L mutant took 72 hours. To rule out any possible effects from the acceptor substrate, the conversion of DL-glyceraldehyde with pyruvate was followed under identical conditions (Figure 8). In this case, no apparent difference was observed between F210L and the WT, demonstrating that the observed effect indeed originated from the different donor substrates.

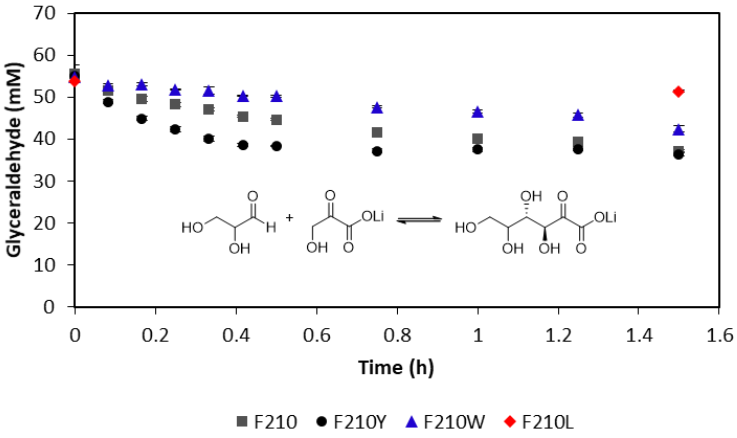


Figure 7: Time course over the first 1.5 hours of the *Sw*HKA catalysed aldol reaction between hydroxypyruvate and DL-glyceraldehyde. Progress was monitored *via* the conversion of glyceraldehyde by HPLC (n = 3). Conditions: *Sw*HKA (0.5 mg/mL), $MgCl_2$ (2 mM), hydroxypyruvate (50 mM), DL-glyceraldehyde (50 mM), KPi buffer (5 mM, pH 7.0, 25°C, 500 rpm).

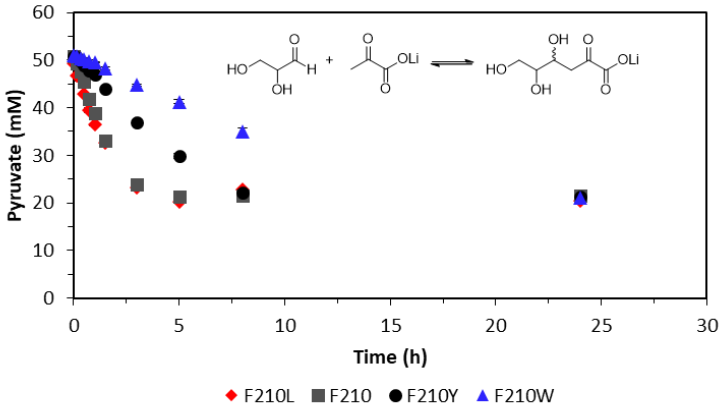


Figure 8: Time course of the *Sw*HKA catalysed aldol reaction between pyruvate and DL-glyceraldehyde. Progress was monitored *via* the conversion of pyruvate by HPLC (n = 3). Conditions: *Sw*HKA (0.5 mg/mL), $MgCl_2$ (2 mM), pyruvate (50 mM), DL-glyceraldehyde (50 mM), KPi buffer (5 mM, pH 7.0, 25°C, 500 rpm).

Soaking of Mg^{2+}-*Sw*HKA crystals with hydroxypyruvate showed that both C-H bonds at C-3 indeed are oriented towards the aromatic system of F210 for CH-π interactions and revealed a hydrogen bond between the substrate's hydroxyl group and the backbone carbonyl of I167 (Figure 2). In this configuration, the hydroxyl group is oriented towards the inside of the active site, preventing it from blocking the narrow substrate channel. Upon deprotonation, the enol intermediate then adopts a planar configuration and docking studies suggest a newly formed hydrogen bond between the hydroxyl group and R69, alongside CH-π interactions of the olefinic hydrogen with F210 (Figure 9).

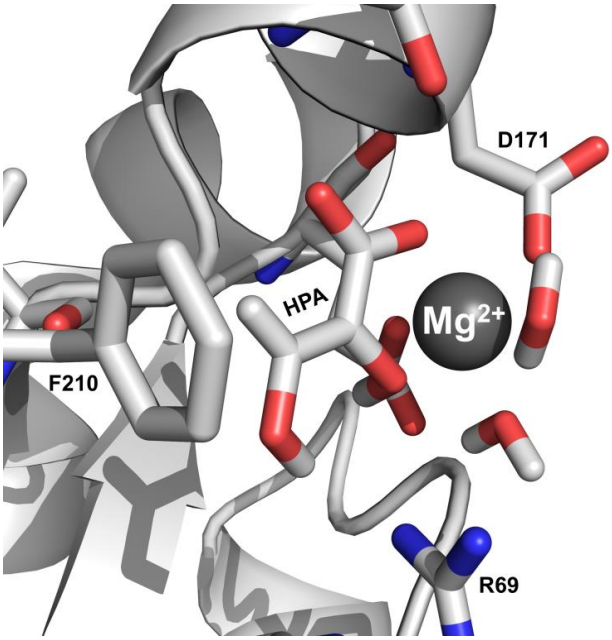


Figure 9: Docking of the planar enol intermediate from HPA into the active site of Mg^{2+}-*Sw*HKA (6r62.pdb) with YASARA. CH-π interactions are established between F210 and the olefinic C-H bond at C_3 in HPA. The hydroxyl group forms a H-bond to R69. For simplicity, the second monomer of *Sw*HKA has been omitted in this representation. Figures were created with PyMOL.[16]

This geometry allows for the acceptor substrate to freely converge towards the nucleophile and facilitates C-C bond formation. In the F210Y mutant, the phenolic hydroxyl group is too far away for direct interactions with the substrate and the observed increase in H-D exchange rate must therefore originate from enhanced CH-π interactions.

Based on the configuration of L216 in the DDG aldolase (1dxf.pdb, rmsd = 1.61 Å), the F210L mutation was modelled *in silico* as the corresponding rotamer. The model revealed steric clashes between the isobutyl side chain and the hydroxyl group, which would prevent both hydroxypyruvate and the enol intermediate from adopting similar configurations. While the F210L mutant can catalyse the first step in the aldol reaction, it subsequently seems to trap the enol intermediate in a disfavoured configuration that hinders C-C bond formation.

4.3 Conclusions

In summary, this characterization of the *Sw*HKA aldolase A5VH82 has yielded new insights into the enzyme's kinetic properties and the mode of phosphate binding and activation. Due to the lack of proximal amino acid residues, a structural molecule of water was previously identified to be responsible for deprotonation of the substrate. However, this water molecule is only insufficiently activated by the adjacent residues and renders deprotonation of the donor substrate rate limiting.

Alternatively, this enzyme can bind inorganic phosphate *via* H44 with high affinity, which replaces water as the catalytic base and results in a rate enhancement of up to 120-fold. Since C-C bond formation becomes rate limiting under phosphate saturated conditions, the overall obtained rate enhancement then depends on the kinetic properties of the acceptor substrate. Our results highlight a common pitfall that is responsible for the previous misconception of the effect of phosphate in class II pyruvate aldolases: excessive phosphate concentrations negatively affect the reaction rate by removing the metal cofactor. We

demonstrated that phosphate activation is not a unique property of the αββα sandwich fold, but also is functionally conserved within the same order of magnitude in the distinct $(\alpha/\beta)_8$ barrel fold of the *Sw*HKA aldolase. Future research into class II pyruvate aldolases should therefore take the possibility of phosphate activation into account.

While the conserved variant F210L efficiently catalyses the first step of the aldol reaction to generate the enol intermediate from HPA, its configuration hinders subsequent C-C bond formation. So far, CH-π interactions have mainly been investigated for carbohydrate-active enzymes (CAZymes). Our findings now illustrate their importance for another enzyme class, where they allow for the conversion of a non-natural substrate. By further increasing the electron density of the aromatic system in the F210Y mutant, we obtained a variant of 2-fold improved activity. CH-π interactions should therefore also be included in rational enzyme design strategies when electropositive hydrogen atoms are involved.

Acknowledgements

Financial support from the ERA-IB scheme (grant ERA-IB-15-110) to L.M. is gratefully acknowledged.

4.4 Methods Summary

Detailed protocols and additional graphs are available in the supplementary information of the *open access* online version of this article.

Oxaloacetate decarboxylase activity assay: *Sw*HKA was incubated with NADH (0.5 mM), M^{2+} (1 mM), L-Lactic dehydrogenase in triethanolamine buffer (TEOA, 5 mM, pH = 7.5) and the reaction was initiated by the addition of oxaloacetate (0.5 mM) to give a final volume of 1 mL. The change in absorbance was followed at 340 nm in PMMA cuvettes (800 rpm, 25°C) in triplicate, using a Cary 60 UV-Vis spectrometer (Agilent Technologies). Curve fitting was performed with IGOR, assuming Michaelis-Menten type kinetics.

H-D exchange assays: NMR spectra were recorded on an Agilent 400 MHz instrument at 298 K with the PRESAT-PURGE pulse sequence. Spectra were recorded using a recycle

delay of 2 seconds and 32 repetitions. A benzene-D_6 capillary was used for external locking and quantitation. *Sw*HKA (1.25 mg/mL) was incubated with Mg^{2+} (2 mM) in potassium phosphate buffer (5 mM, pD = 7.4, prepared in D_2O, lyophilised and re-dissolved in D_2O). The reaction was initialised by the addition of ketoacid substrate (50 mM). Controls in the absence of enzyme under identical conditions showed <1% chemical background activity over the investigated time frame.

Aldol reactions of (hydroxy-) pyruvate with DL-glyceraldehyde: *Sw*HKA (0.5 mg/mL) was incubated with Mg^{2+} (2 mM), DL-glyceraldehyde (50 mM) and the reaction was initialised by the addition of ketoacid substrate (50 mM) in potassium phosphate buffer (5 mM, pH 7.0, 1.5 mL, 25°C, 500 rpm). Samples were quenched by 1:1 dilution with trifluoroacetic acid (TFA, 0.2% v/v in H_2O) and the enzyme was precipitated by centrifugation. Concentrations were determined by RP-HPLC (ICSep ICE Coregel-87H3 column, 0.4x25 cm, Transgenomic, 0.1% v/v TFA, 0.8 mL/min, 60°C, 210 nm) on a Shimadzu LC-20AD system using external standards.

Crystallisation and data collection: Well diffracting crystals of apo-*Sw*HKA protein were obtained at 277 K using the vapour diffusion method. Drops were made by mixing equal volumes of a protein sample at 9.2 mg/ml with a crystallisation solution consisting of 0.44 M to 0.65 M of sodium citrate and 0.1 M HEPES (4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid) at pH 7.05. Crystals took 7days to appear and grew to their final size within a few days. Crystals of the holo-*Sw*HKA in complex with hydroxypyruvate were obtained by transferring crystals of apo-*Sw*HKA into a new drop containing 0.7 M of potassium bromide (KBr), 0.1 M HEPES pH 7.05, 20 mM of magnesium chloride ($MgCl_2$) and 150 mM of hydroxypyruvate. The crystals were left in this new condition for a few hours and then were cryo-cooled using as cryoprotectant 20% of glycerol. X-ray diffraction data was collected at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany).[32] The protein crystallized in the cubic space group P 42 3 2, with unit-cell parameters of approx. a = 116.6 Å, with one molecule in the asymmetric unit and a solvent content of 47.5%.

Structure solution and refinement: Data were indexed and integrated with XDS,[33] scaled with AIMLESS,[34] and the space group was determined with POINTLESS[35] from CCP4 programs suite.[36] The structure was solved by the molecular-replacement method with the program MOLREP[37] using the crystal structure of apo-*Sw*HKA as a search model. Model building and refinement of ligand bound holo-*Sw*HKA was performed with COOT[38] and PHENIX,[39] respectively. The program PHENIX was used to refine atomic coordinates together with individual isotropic atomic displacement parameters. TLS thermal anisotropic parameterization was also included in the final stages of refinement, with each monomeric subunit divided into three TLS groups, as suggested by TLS Motion Determination.[38, 40] A randomly selected 5% of observed reflections were kept aside during refinement for cross-validation. The positions of the magnesium ion and of the hydroxypyruvate became evident

after the first cycles of refinement, based on electron density difference maps, and both entities were modelled and refined with full occupancy. Solvent molecules were included in the model after a few rounds of refinement as well as bromide and potassium ions that were located at the surface of the protein.

Substrate docking: A simulation cell of 9 Å was defined around R69 in the crystal structure of Mg^{2+}-*Sw*HKA (6r62.pdb). The enol intermediate (2,3-dihydroxy acrylic acid) was energy minimised (MM2 calculation, Chem3D 15.1) and subsequently docked into the active site using YASARA (dock_run.mcr, version 16.12.29). A structural model for variant F210L was created *in silico* based on the rotamer orientation of L216 in the DDG pyruvate aldolase (1dxf.pdb) and the enol intermediate was docked into the model following the same approach.

References

[1] a) K. Hernández, J. Joglar, J. Bujons, T. Parella, P. Clapés, *Angew. Chem. Int. Ed.* **2018**, *57*, 3583-3587; b) K. Hernandez, J. Bujons, J. Joglar, S. J. Charnock, P. Dominguez de Maria, W. D. Fessner, P. Clapés, *ACS Catal.* **2017**, *7*, 1707-1711.

[2] V. Laurent, E. Darii, A. Aujon, M. Debacker, J. L. Petit, V. Hélaine, T. Liptaj, M. Breza, A. Mariage, L. Nauton, *Angew. Chem.* **2018**, *130*, 5565-5569.

[3] a) W. A. Greenberg, A. Varvak, S. R. Hanson, K. Wong, H. Huang, P. Chen, M. J. Burk, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5788-5793; b) J. Wagner, R. A. Lerner, C. F. Barbas, *Science* **1995**, *270*, 1797-1800; c) C. F. Barbas, A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Björnestedt, B. List, J. Anderson, E. A. Stura, I. A. Wilson, *Science* **1997**, *278*, 2085-2092; d) R. N. Patel, in *Organic Synthesis Using Biocatalysis*, Elsevier, **2015**, pp. 339-411.

[4] A. Liese, K. Seelbach, A. Buchholz, J. Haberland, in *Industrial Biotransformations*, second ed. (Eds.: A. Liese, K. Seelbach, C. Wandrey), Wiley-VCH, **2006**, p. 459.

[5] a) C.-H. Wong, G. M. Whitesides, *Enzymes in synthetic organic chemistry, Vol. 12*, Academic Press, **1994**; b) A. Bolt, A. Berry, A. Nelson, *Arch. Biochem. Biophys.* **2008**, *474*, 318-330; c) P. Clapés, in *Science of Synthesis Biocatalysis in Organic Synthesis 2* (Ed.: W.-D. F. K. Faber, N.J. Turner), Georg Thieme Verlag, Stuttgart, **2015**, pp. 31-93; d) J. Sukumaran, U. Hanefeld, *Chem. Soc. Rev.* **2005**, *34*, 530-542; e) A. Ranoux, U. Hanefeld, in *Stereoselective Synthesis of Drugs and Natural Products*, First ed. (Ed.: V. Andrushko, N. Andrushko), John Wiley & Sons, **2013**, pp. 831-857.

[6] B. Horecker, O. Tsolas, C. Lai, in *The enzymes, Vol. 7*, Elsevier, **1972**, pp. 213-258.

[7] a) D. E. Morse, B. Horecker, *Advances in Enzymology and Related Areas of Molecular Biology, Volume 31* **1968**, 125-181; b) W. Rutter, in *Fed. Proc., Vol. 23*, **1964**, pp. 1248-1257.

[8] a) K. Furukawa, J. Hirose, A. Suyama, T. Zaiki, S. Hayashida, *J. Bacteriol.* **1993**, *175*, 5224-5232; b) K. Furukawa, N. Kimura, *Environ. Health Perspect.* **1995**, *103*, 21-23; c) J.-S. Seo, Y.-S. Keum, Y. Hu, S.-E. Lee, Q. X. Li, *Biodegradation* **2007**, *18*, 123-131; d) E. Masai, Y. Katayama, M. Fukuda, *Biosci., Biotechnol., Biochem.* **2007**, *71*, 1-15.

[9] a) P. Baker, S. Y. Seah, *J. Am. Chem. Soc.* **2011**, *134*, 507-513; b) M. Cheriyan, M. J. Walters, B. D. Kang, L. L. Anzaldi, E. J. Toone, C. A. Fierke, *Biorg. Med. Chem.* **2011**, *19*, 6447-6453; c) M. J. Walters, E. J. Toone, *Nat. Protoc.* **2007**, *2*, 1825-1830.

[10] V. De Berardinis, C. Guérard-Hélaine, E. Darii, K. Bastard, V. Hélaine, A. Mariage, J.-L. Petit, N. Poupard, I. Sánchez-Moreno, M. Stam, *Green Chem.* **2017**, *19*, 519-526.

[11] T. Izard, N. C. Blackwell, *EMBO J.* **2000**, *19*, 3849-3856.

[12] W. Wang, S. Y. Seah, *Biochemistry* **2005**, *44*, 9447-9455.

[13] W. Wang, S. Mazurkewich, M. S. Kimber, S. Y. Seah, *J. Biol. Chem.* **2010**, *285*, 36608-36615.

[14] S. Mazurkewich, S. Y. Seah, *PLoS ONE* **2016**, *11*, e0164556.

[15] M. Coincon, W. Wang, J. Sygusch, S. Y. Seah, *J. Biol. Chem.* **2012**, *287*, 36208-36221.

[16] W. L. DeLano, *DeLano Scientific, San Carlos, CA* **2002**, *452*.

[17] K. L. Hudson, G. J. Bartlett, R. C. Diehl, J. Agirre, T. Gallagher, L. L. Kiessling, D. N. Woolfson, *J. Am. Chem. Soc.* **2015**, *137*, 15152-15160.

[18] a) A. C. Joerger, C. Gosse, W.-D. Fessner, G. E. Schulz, *Biochemistry* **2000**, *39*, 6033-6041; b) S. J. Cooper, G. A. Leonard, S. M. McSweeney, A. W. Thompson, J. H. Naismith, S. Qamar, A. Plater, A. Berry, W. N. Hunter, *Structure* **1996**, *4*, 1303-1315; c) M. Kroemer, I. Merkel, G. E. Schulz, *Biochemistry* **2003**, *42*, 10560-10568.

[19] a) S. Kobayashi, S. Nagayama, T. Busujima, *J. Am. Chem. Soc.* **1998**, *120*, 8287-8288; b) S. Kobayashi, K. Manabe, *Acc. Chem. Res.* **2002**, *35*, 209-217.

[20] F. Vertregt, G. Torrelo, S. Trunk, H. Wiltsche, W. R. Hagen, U. Hanefeld, K. Steiner, *ACS Catalysis* **2016**, *6*, 5081-5085.

[21] O. C. Richards, W. J. Rutter, *J. Biol. Chem* **1961**, *236*, 3177-3184.

[22] a) W. Wang, S. Y. Seah, *FEBS Lett.* **2008**, *582*, 3385-3388; b) P. Baker, J. Carere, S. Y. Seah, *Biochemistry* **2011**, *50*, 3559-3569; c) D. Rea, R. Hovington, J. F. Rakus, J. A. Gerlt, V. Fülöp, T. D. Bugg, D. I. Roper, *Biochemistry* **2008**, *47*, 9955-9965; d) D. Rea, V. Fülöp, T. D. Bugg, D. I. Roper, *J. Mol. Biol.* **2007**, *373*, 866-876; e) G.-T. Huang, J.-S. K. Yu, *ACS Catal.* **2017**, *7*, 8130-8133.

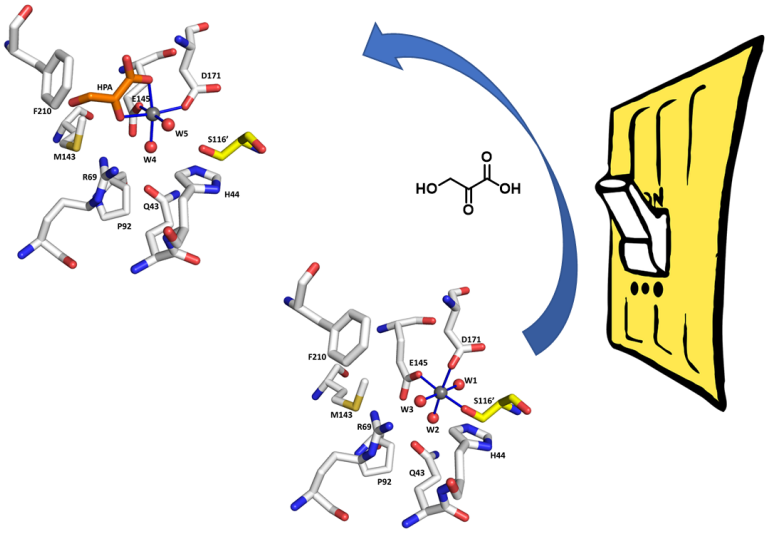
[23] a) N. Chickerur, G. Nayak, R. Lenka, P. Mahapatra, *Thermochim. Acta* **1982**, *58*, 111-115; b) N. E. Good, G. D. Winget, W. Winter, T. N. Connolly, S. Izawa, R. M. Singh, *Biochemistry* **1966**, *5*, 467-477.

[24] M. E. Smith, K. Smithies, T. Senussi, P. A. Dalby, H. C. Hailes, *Eur. J. Org. Chem.* **2006**, *2006*, 1121-1123.

[25] S. R. Marsden, L. Gjonaj, S. J. Eustace, U. Hanefeld, *ChemCatChem* **2017**, *9*, 1808-1814.
[26] A. K. Whiting, Y. R. Boldt, M. P. Hendrich, L. P. Wackett, L. Que, *Biochemistry* **1996**, *35*, 160-170.
[27] a) A. K. Whiting, Y. R. Boldt, M. P. Hendrich, L. P. Wackett, L. Que, *Biochemistry* **1996**, *35*, 160-170; b) G. H. Reed, S. D. Morgan, *Biochemistry* **1974**, *13*, 3537-3541.
[28] P. Campomanes, W. F. Kellett, L. M. Easthon, A. Ozarowski, K. N. Allen, A. Angerhofer, U. Rothlisberger, N. G. Richards, *J. Am. Chem. Soc.* **2014**, *136*, 2313-2323.
[29] S. W. Fullerton, J. S. Griffiths, A. B. Merkel, M. Cheriyan, N. J. Wymer, M. J. Hutchins, C. A. Fierke, E. J. Toone, J. H. Naismith, *Biorg. Med. Chem.* **2006**, *14*, 3002-3010.
[30] a) W. Chen, S. Enck, J. L. Price, D. L. Powers, E. T. Powers, C.-H. Wong, H. J. Dyson, J. W. Kelly, *J. Am. Chem. Soc.* **2013**, *135*, 9877-9884; b) M. d. C. Fernández-Alonso, F. J. Cañada, J. Jiménez-Barbero, G. Cuevas, *J. Am. Chem. Soc.* **2005**, *127*, 7379-7386; c) J. L. Asensio, A. Ardá, F. J. Cañada, J. s. Jiménez-Barbero, *Acc. Chem. Res.* **2012**, *46*, 946-954; d) M. I. Chávez, C. Andreu, P. Vidal, N. Aboitiz, F. Freire, P. Groves, J. L. Asensio, G. Asensio, M. Muraki, F. J. Cañada, *Chem. Eur. J.* **2005**, *11*, 7060-7074.
[31] a) I. Salmeen, G. Palmer, *J. Chem. Phys.* **1968**, *48*, 2049-2052; b) A. Lundin, R. Aasa, *J. Magn. Reson.* **1972**, *8*, 70-73.
[32] M. Cianci, G. Bourenkov, G. Pompidor, I. Karpics, J. Kallio, I. Bento, M. Roessle, F. Cipriani, S. Fiedler, T. R. Schneider, *J. Synchrotron Radiat.* **2017**, *24*, 323-332.
[33] W. Kabsch, *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010**, *66*, 125-132.
[34] P. R. Evans, *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2011**, *67*, 282-292.
[35] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2013**, *69*, 1204-1214.
[36] C. P. Collaborative, *Acta Crystallogr. D Biol. Crystallogr.* **1994**, *50*, 760.
[37] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022-1025.
[38] J. Painter, E. A. Merritt, *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2005**, *61*, 465-471.
[39] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010**, *66*, 213-221.

5

Substrate Induced Movement of the Metal Cofactor between Active and Resting State Configurations in a Metalloenzyme



This chapter is based on:

Stefan R. Marsden, Hein J. Wijma, Michael Mohr, Peter-Leon Hagedoorn, Jesper Laustsen, Cy M. Jeffries, Luuk Mestrom, Duncan G. G. McMillan, Isabel Bento and Ulf Hanefeld

Manuscript in preparation

Regulation of enzyme activity is vital for living organisms. In metalloenzymes, far-reaching rearrangements of the protein scaffold are generally required to tune the metal cofactor's properties by allosteric regulation. Here, we determined several crystal structures of the metal containing hydroxyketoacid aldolase from *Sphingomonas wittichii* RW1 (*Sw*HKA) up to a resolution of 1.2 Å. This structural analysis revealed a dynamic movement of the metal cofactor between two coordination spheres of distinct properties in response to changes within its external environment. In its resting state configuration (M^{2+}_R), the metal constitutes an integral part of the dimer interface within the overall hexameric assembly, but sterical constraints do not allow for substrate binding in this position. Conversely, an alternative confirmation, in which the metal has moved by 2.4 Å constitutes the catalytically active state (M^{2+}_A). Bidentate coordination of a ketoacid substrate to M^{2+}_A affords the overall lowest energy complex, which drives the equilibrium transition from M^{2+}_R to M^{2+}_A without a concomitant change of protein structure. Our results demonstrate how modulation of the energy of formation for different protein-metal complexes by dynamic ligand exchange allows for the environment responsive regulation of enzyme properties without a concomitant rearrangement of the protein structure.

5.1 Introduction

An estimated one third of all known enzymes require a metal cofactor to display their biological function.[1-4] The formed protein-metal complexes are typically well-defined both in terms of coordination geometry, and their exact position within the respective metalloenzymes.[5,6] In order to allow for a dynamic adaption to changed external conditions, the physicochemical properties of the metal center are commonly controlled by allosteric regulation.[7] In this mechanism, far-reaching rearrangements of the protein structure are generally required to tune the metal cofactor's properties *via* discrete movements of the coordinating residues.[8-10]

Alternatively, a dynamic transition of the metal cofactor between distinct coordination sites could similarly alter its physicochemical properties without the need for a concomitant movement of protein residues. Such a dynamic behavior of the metal cofactor was previously reported to constitute the key element in the catalytic mechanism of D-xylose isomerase[11] and in members of the medium chain alcohol dehydrogenase superfamily (MDR).[12,13] In the MDR superfamily, zinc acts as a Lewis acid for the activation of aldehydes and ketones, which makes them susceptible to reduction by NAD(P)H.[12] Conversely, the enzyme facilitates an increase in electron density for the oxidation of alcohols by $NAD(P)^+$ in reverse. Depending on the direction of the reaction, the metal cofactor must therefore exert either an electron donating-, or an electron withdrawing effect.[12] This is achieved by variation between two coordination sites, which differently influence the protonation state of a coordinated water molecule and facilitate proton transfer in the required direction.[12]

Similarly, D-xylose isomerase catalyzes the isomerization of D-glucose to D-fructose *via* the dynamic movement of one of its two magnesium cofactors. M1 remains stationary throughout the catalytic cycle and is responsible for substrate binding. Conversely, M2 alternates between two binding sites to facilitate the

sequential activation of the C-1 aldehyde (electron withdrawing effect for reduction at position M2a) and activation of the C-2 hydroxyl group (electron donating effect for oxidation at position M2b) and *vice versa*. This dynamic transition between M2a and M2b facilitates the redox neutral hydride transfer for isomerization of the corresponding aldose and ketose forms.[11] In both examples, the dynamic behavior of the metal cofactor constitutes an essential aspect of the enzymes' catalytic mechanisms.

Here, we describe the substrate induced movement of the metal cofactor between a catalytically active and inactive coordination sphere within the divalent metal containing hydroxy ketoacid aldolase from *Sphingomonas wittichii* RW1 (*Sw*HKA). This enzyme displays an unusual promiscuity toward structural analogues of pyruvate,[14] which were attributed to CH-π interactions with aromatic amino acids that replaced a conserved leucine residue.[15] Notably, *Sw*HKA is activated by trace amounts of inorganic phosphate ($K_{d,Pi}$ = 175 µM), which effect a rate enhancement of up to two orders in magnitude for the aldol reaction.[15] Its synthetic potential has been demonstrated for the conversion of various pyruvate analogues, including hydroxypyruvate,[14,15] fluoropyruvate[16] and 2-oxobutyrate.[17]

Based on these reports, we conducted a detailed structural characterization of *Sw*HKA, which revealed two coordination sites of distinct properties. In contrast to allosteric regulation, the structural changes between the enzyme's active (M^{2+}_A) and resting state (M^{2+}_R) configurations are confined to the position of the metal cofactor. Most importantly, *Sw*HKA therefore seems to be selectively activated by the characteristic structural motif of its ketoacid substrates.

5.2 Results

Crystal structures

*Sw*HKA (accession number A5VH82, EC 4.1.2.52) was recombinantly expressed in *E. coli* BL21(DE3) and subsequently purified *via* its *N*-terminal His_6-tag, followed by size-exclusion chromatography to afford the *apo*-enzyme in high purity, as described previously.[15] The *de novo* crystal structure of wild-type *apo*-*Sw*HKA was solved by single-wavelength anomalous dispersion of the sulfur atoms (S-SAD), originating from two cysteine and six methionine residues per monomer. This structure was subsequently used as search model for solving the crystal structures of wild-type and mutant variants of *Sw*HKA in the presence of different metals and substrates up to 1.2 Å resolution.

Holo-*Sw*HKA forms a trimer of homodimers, in which each monomer adopts an $(\alpha/\beta)_8$ triosephosphate isomerase (TIM) barrel fold (Fig. 1). Domain swapping of the 8th α-helix packing onto a β-sheet affords a dimer, and trimerization thereof generates the final hexameric assembly. In total, *holo*-*Sw*HKA contains six equivalent active sites; each of which are formed from residues from two different domain-swapped dimers. This makes *Sw*HKA a functional hexamer. Notably, formation of the overall hexameric assembly of *apo*-*Sw*HKA does not require the presence of a metal cofactor, as was evident from small angle x-ray scattering.

Active and resting state configurations in *Sw*HKA

Crystals of WT *apo*-*Sw*HKA and mutant variants thereof were soaked with $MgCl_2$ or $MnCl_2$ to afford the *holo*-enzyme. Under these conditions, the metal complex was found to be formed at the interface between two domain swapped dimers (Fig 1d). Here, the metal is coordinated by residues E145 and D171 in a monodentate fashion, next to S116’ (where the prime symbol denotes residues from a second monomeric subunit), and three structural water molecules.

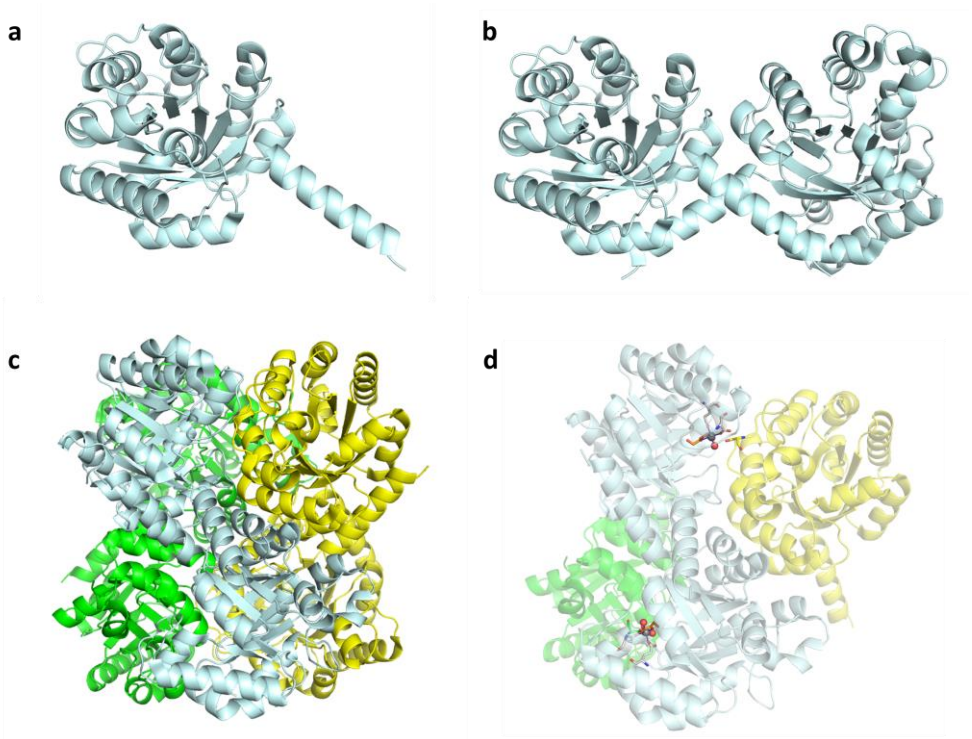


Figure 1: Crystal structure of *Sw*HKA. **(a)** TIM barrel fold monomeric *Sw*HKA. **(b)** Two α-helices pack onto β-sheets to afford a dimer *via* domain swapping (both monomers in cyan). **(c)** Trimerization of domain swapped dimers (green, yellow, cyan) affords the overall hexameric holoenzyme. **(d)** Partial representation of *holo*-*Sw*HKA, showing the location of active sites as ball and stick models at the interfaces of two domain swapped dimers. The second monomers (green, yellow) were omitted for visualization purposes.

In this configuration, replacement of either one of the three bound water molecules for substrate coordination is sterically impossible due to structural clashes with residues Q43, R69 and P92 (Fig. 2a). Notably, the bidentate coordination of ketoacids was previously shown to be essential for catalytic

activity.[15] The coordination sphere S116', E145 and D171 therefore constitutes a catalytically inactive resting state of the enzyme (M^{2+}_{R}).

Conversely, a square pyramidal coordination sphere with the metal at 2.4 Å distance from that of M^{2+}_{R} constitutes the catalytically active state of *holo-Sw*HKA. Here, the coordination sphere no longer includes residue S116', but only consists of the five ligands E145, D171 next to three water molecules ($M^{2+}_{A,V}$, Fig. 2b). The replacement of S116' by a water ligand moves the metal cofactor toward the entrance of the active site, while the involved protein side chains retain their original position. This localized movement of the metal cofactor generates a vacant coordination site ($M^{2+}_{A,V}$) and newly creates sufficient space for bidentate substrate binding *via* the exchange of the water ligand W6 (Fig. 2c). In the absence of substrates, M^{2+}_{R} constitutes the predominant state and the $M^{2+}_{A,V}$ complex was observed with lower occupancy in crystal structures of *holo-Sw*HKA. Assuming a dynamic equilibrium between both coordination spheres, M^{2+}_{R} must therefore be lower in energy than $M^{2+}_{A,V}$.

When the crystals were soaked with ketoacids, bidentate substrate coordination was observed as the predominant state by formation of the lowest energy complex ($M^{2+}_{A,S}$, Fig. 2c). The $M^{2+}_{A,S}$ complex is stabilized not only by the oxygen-metal interactions, but also by side chain interactions from the protein environment with the coordinated substrate; e.g. CH-π interactions between the electropositive hydrogen atoms at C-3 in hydroxypyruvate (HPA) and the aromatic residue F210.[15] Bidentate substrate coordination therefore seems to constitutes the driving force for the transition of the metal cofactor from M^{2+}_{R}, probably *via* $M^{2+}_{A,V}$ into the catalytically active configuration $M^{2+}_{A,S}$.

A stoichiometry of one metal ion per active site was determined by isothermal calorimetry titrations (ITC). Next to sterical constraints, this observation excludes the simultaneous binding of two metal ions to both M^{2+}_{R} and M^{2+}_{A} in *Sw*HKA.

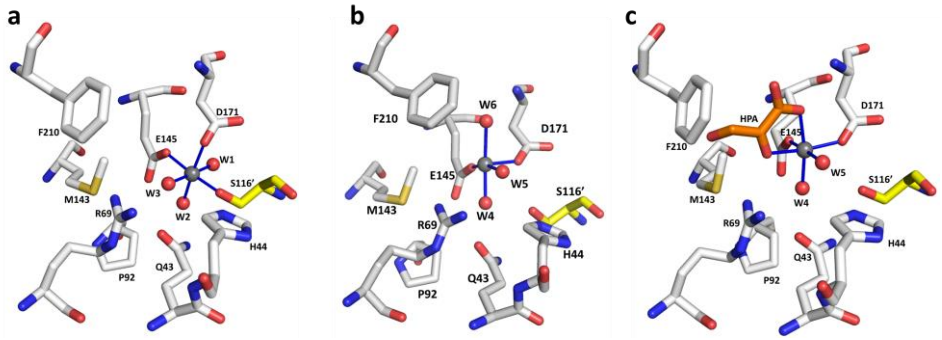
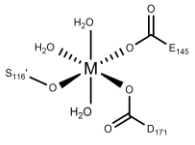


Figure 2: Active and resting state coordination spheres in *Sw*HKA. **(a)** M^{2+}_{R} coordination sphere of the metal cofactor at the interface between two dimers (white; yellow, denoted by the prime symbol). M^{2+}_{R} was observed in the following crystals structures: WT-*Sw*HKA-Mg; WT-*Sw*HKA-Mn; F210W-*Sw*HKA-Mg; H44A-*Sw*HKA-Mg. **(b)** Square pyramidal $M^{2+}_{A,V}$ coordination sphere with a vacant coordination site. This complex was observed with low occupancy in the absence of substrates in the crystal structure of WT-*Sw*HKA-Mn. **(c)** Coordination sphere of the catalytically active state $M^{2+}_{A,S}$. This complex was observed in the presence of substrate (hydroxypyruvate) in WT-*Sw*HKA-HPA (6r62.pdb) and F210W-*Sw*HKA-HPA.

QM energy calculations of metal complexes

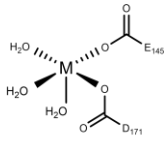
Our findings suggest that the three coordination states of the metal cofactor M^{2+}_{R}, $M^{2+}_{A,V}$ and $M^{2+}_{A,S}$ are part of a dynamic equilibrium, in which their relative distribution is determined by the energy difference between the corresponding metal complexes. We therefore set out to calculate the energies of formation using quantum mechanics (QM). The coordinating protein residues were approximated as acetate (for E145 and D171) and methanol (for S116') molecules within an aqueous environment, which is common for a QM-only approach.[16,18,19] The predicted geometries of the metal complexes were found in good agreement with the experimentally determined structures. While accurate predictions of the equilibrium distributions between M^{2+}_{R}, $M^{2+}_{A,V}$ and $M^{2+}_{A,S}$ would require an extensive computational treatment of the whole active site environment, the

general trend was in good agreement with experimental observations (Fig. 3). The square-pyramidal $M^{2+}_{A,V}$ complex was unanimously predicted for both Mg^{2+} and high-spin Mn^{2+} (the observed species in *Sw*HKA[15]) to be lower in energy than the corresponding hypothetical octahedral $M^{2+}_{A,V}$ complex with four water ligands. Thermodynamics therefore favor a square-pyramidal coordination geometry for $M^{2+}_{A,V}$, which promotes substrate binding by providing a free coordination site. Similarly, $M^{2+}_{A,S}$ was unanimously predicted to constitute the overall lowest energy complex. This confirmed our previous notion, that bidentate substrate coordination constitutes the driving force for the enzyme's transition from the resting state into its catalytically active form.



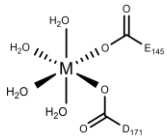
1

	APFD (kJ/mol)	B3LYP (kJ/mol)
Mg^{2+}	0.0	0.0
Mn^{2+}	0.0	0.0



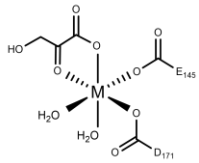
2

	APFD (kJ/mol)	B3LYP (kJ/mol)
Mg^{2+}	-11.7	-12.0
Mn^{2+}	18.5	-19.9



3

	APFD (kJ/mol)	B3LYP (kJ/mol)
Mg^{2+}	2.6	25.9
Mn^{2+}	38.0	24.8



4

	APFD (kJ/mol)	B3LYP (kJ/mol)
Mg^{2+}	-58.5	-44.8
Mn^{2+}	-48.6	-35.8

Figure 3: Metal complexes used for QM-only calculations. Acetate and methanol ligands were used to approximate protein residues. Complex **1** represents M^{2+}_{R}, complex **2** represents $M^{2+}_{A,V}$. Complex **3** represents a hypothetical $M^{2+}_{A,V}$ complex with four water ligands in octahedral configuration, and complex **4** represents $M^{2+}_{A,S}$. Two different force fields (APFD and B3LYP) were used for the calculation of relative energies of formation, which were normalized for the resting state M^{2+}_{R}. For complex **2**, only the calculations for high-spin Mn^{2+} with APFD matched with experimental observations of M^{2+}_{R} being lower in energy than square-pyramidal $M^{2+}_{A,V}$.

Structural alignments

In a previous study, several variants of *Sw*HKA were created at position F210 in order to probe the influence of aromatic electron density for CH-π interactions with hydroxypyruvate.[15] Here, the crystal structure of variant F210W soaked with $MgCl_2$ and hydroxypyruvate could now be solved up to 1.2 Å resolution. While WT *Sw*HKA showed a monomer in the asymmetric unit with an occupancy of 1.0 for the metal cofactor in $M^{2+}_{A,S}$ and coordinated hydroxypyruvate (6r62.pdb), variant F210W crystallized in a different space group and a dimer was observed in the asymmetric unit. In this F210W crystal structure, monomer A exclusively contained the metal cofactor in the M^{2+}_{R} configuration with an occupancy of 1.0, whereas monomer B displayed an occupancy of 0.7 for $M^{2+}_{A,S}$ and 0.3 for the catalytically inactive M^{2+}_{R}. Both monomers are structurally highly similar (rmsd = 0.168 Å) and differ solely with respect to the position of a few solvent exposed, mobile loops. Alignment of WT *Sw*HKA $M^{2+}_{A,S}$ (6r62.pdb) with monomer B from variant F210W (containing $M^{2+}_{A,S}$/M^{2+}_{R}) showed no significant change in protein structure by the F210W mutation (Fig. S1, rmsd = 0.194 Å). Possible communication in the form allosteric regulation, that would render the two active sites of the homodimer chemically non-equivalent could therefore not be identified. It is assumed that diffusion limitations reduced the availability of substrate during the soaking of larger F210W crystals with hydroxypyruvate and led to this snapshot of a (yet) incomplete transition from M^{2+}_{R} to $M^{2+}_{A,S}$.

*Sw*HKA displays activity both with Mn^{2+} and Mg^{2+}. Soaking of WT *apo*-*Sw*HKA crystals with both Mg^{2+} or Mn^{2+} in the absence of substrates did not result in any change of protein structure (Fig. S2, rmsd = 0.164 Å), but led to different distributions between M^{2+}_{R} and $M^{2+}_{A,V}$. With Mg^{2+}, the M^{2+}_{R} configuration showed an occupancy of 1.0, whereas a mixture of M^{2+}_{R} (0.57) and $M^{2+}_{A,V}$ (0.43) was observed for Mn^{2+}. In this case, the complete saturation of both active sites with the metal cofactors suggests that the observed relative occupancy of each state

accurately reflects its actual equilibrium distribution. The energy difference between M^{2+}_{R} and $M^{2+}_{A,V}$ in *Sw*HKA is therefore substantially larger for protein-metal complexes with Mg^{2+} than with Mn^{2+}.

Thermal stability of *Sw*HKA variants

To investigate a potentially structural role for the metal cofactor in the M^{2+}_{R} configuration through interactions with residue S116', two different variants were designed at this position. In the first variant, serine was replaced by alanine (S116A) to circumvent the formation of M^{2+}_{R}. Conversely, a cysteine mutation (S116C) was introduced to potentially alter the binding in M^{2+}_{R}. Notably, mutagenesis at position S116 reduced the expression yield of soluble protein by more than 10-fold down to ≈3 mg/L of expression medium in both cases. Unfortunately, no well-diffracting crystals could be obtained for both variants despite continued efforts, which limited our analysis to the interpretation of biochemical data. Being an integral part of the dimer interface, M^{2+}_{R} was expected to increase the thermal stability of wild-type holo-*Sw*HKA. We therefore determined protein melting temperatures using differential scanning calorimetry (DSC, Fig. 4). In the absence of metals, wild-type *apo*-*Sw*HKA and *apo*-S116A displayed similar melting temperatures (T_m) of approximately 66 °C, while variant S116C showed a slightly lower stability of 62 °C. Both variants therefore did not seem to substantially disrupt the protein fold. Surprisingly, no significant difference was observed between the wild-type and S116A variants of *Sw*HKA upon holoenzyme formation with $MgCl_2$, as the thermal stability of all variants increased by ≈7°C. This suggests that the thermal unfolding of *Sw*HKA is not initiated by the dissociation of the domain swapped dimers from each other, which prevents measurements of the potential increase in stability by M^{2+}_{R}. In the presence of hydroxypyruvate, all three variants displayed the same melting temperature of 73°C. This is in line with expectations, as the $M^{2+}_{A,S}$ complex should be identical in all three variants.

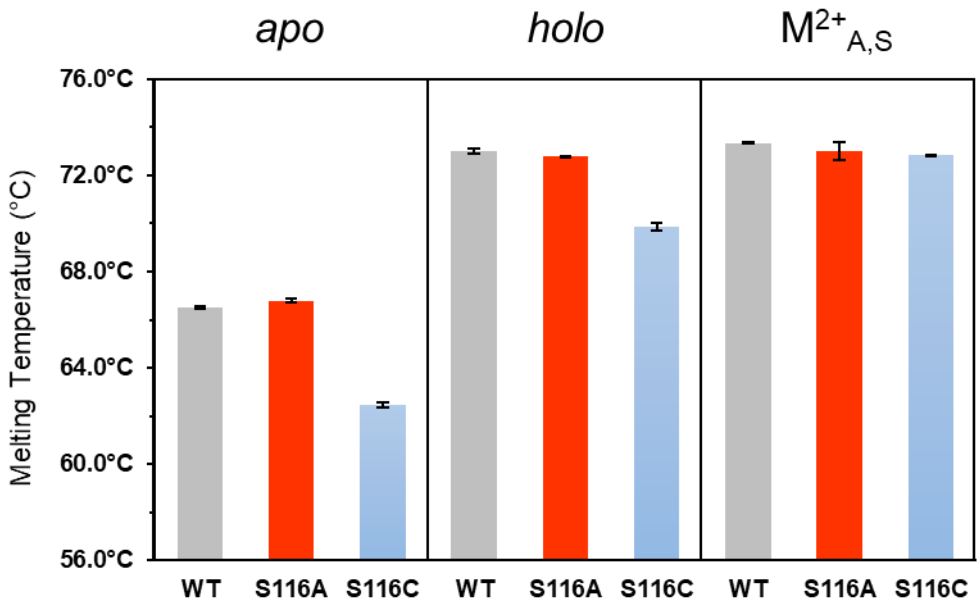


Figure 4: Protein melting temperatures. Thermal unfolding was investigated by differential scanning calorimetry for variants of *Sw*HKA under different conditions. The holoenzyme state could correspond to both M^{2+}_{R} and $M^{2+}_{A,V}$. Conditions: *Sw*HKA (5 mg/mL), $MgCl_2$ (10 mM), hydroxypyruvate (20mM), in HEPES buffer (20 mM, pH 7.5).

The Mn^{2+} cofactor in WT *Sw*HKA was previously analyzed by electron paramagnetic resonance (EPR), and showed an unusually broad signal[15] due to a distorted octahedral geometry.[20] Analysis of the Mn^{2+} complexes of *holo*-*Sw*HKA variants S116A and S116C in the absence of hydroxypyruvate showed largely identical EPR signals. The cysteine residue in S116C therefore does not seem to be involved in the coordination of the metal cofactor in M^{2+}_{R}. In the presence of hydroxypyruvate, both variants similarly showed a broad EPR signal.

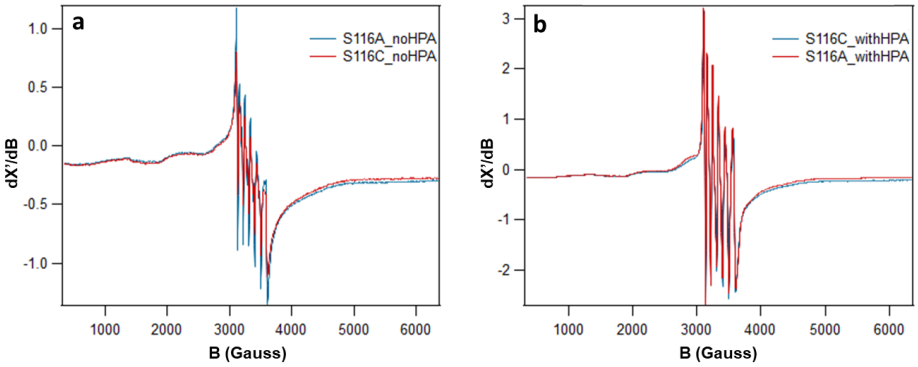


Figure 5: EPR spectrums of *Sw*HKA variants S116A and S116C in the presence/absence of coordinating hydroxypyruvate. Conditions: 1.2 mM *apo-Sw*HKA, 1 mM Mn^{2+} (0.85 eq.) 2 mM KPi, hydroxypyruvate (10 mM) in 5 mM TEOA, pH 7.5, 9.402 GHz microwave frequency, 0.2 mW microwave power, 100 kHz modulation frequency, 10 Gauss modulation amplitude, 37 K.

Metal dissociation constants

The affinity of the different enzyme variants towards Mg^{2+} and Mn^{2+} as cofactors were investigated both in the presence and absence of phosphate as a kinetic activator for *Sw*HKA. Notably, the direct coordination of phosphate to the metal was previously excluded by EPR measurements.[15] For this purpose, retro-aldol activity was used as a reporter to determine the apparent constants K_d and v_{sat} for the $M^{2+}_{A,S}$ complex in metal saturation experiments (Table 1). For the S116A mutant, a lower apparent K_d (i.e. higher affinity) toward $M^{2+}_{A,S}$ was observed with Mg^{2+}, while the affinity remained mostly unaffected when Mn^{2+} was used. These data suggest that elimination of M^{2+}_{R} could remove the competition between the two binding modes, and thereby increase its affinity toward the $M^{2+}_{A,S}$ configuration. Slightly higher apparent metal dissociation constants (i.e. lower affinity toward $M^{2+}_{A,S}$) were observed with variant S116C. This could be due to unfavorable sterical effects, which are difficult to rationalize in the absence of a crystal structure, but would be in line with its lower melting temperature (Fig. 4).

Table 1: Overview of metal dissociation constants (K_d) and maximum rates (v_{sat}) for *Sw*HKA variants under different conditions. Assays were carried out in the presence and absence of inorganic phosphate as kinetic activator, based on the retro-aldol type cleavage of oxaloacetate into pyruvate.[15] Since activity was used as reporter, the data shown are with respect to the $M^{2+}{}_{A,S}$ complex.

Variant	M^{2+}	Phosphate	v_{sat} (U/mg)	K_d (µM)
WT[15]	Mg^{2+}	-	1.65±0.03	58.4±4.1
S116A	Mg^{2+}	-	3.66±0.05	37.9±3.0
S116C	Mg^{2+}	-	2.78±0.12	80.4±19.4
WT[15]	Mg^{2+}	+	17.21±0.29	158±12
S116A	Mg^{2+}	+	9.27±0.27	105±18
S116C	Mg^{2+}	+	5.23±0.1	179±16
WT[15]	Mn^{2+}	-	2.15±0.09	3.3±0.6
S116A	Mn^{2+}	-	3.2±0.09	3.4±0.6
S116C	Mn^{2+}	-	2.3±0.06	4.1±0.6
WT[15]	Mn^{2+}	+	27.0±0.67	6.5±0.7
S116A	Mn^{2+}	+	8.72±0.22	9.1±1.3
S116C	Mn^{2+}	+	4.28±0.16	26.0±5.2

5.3 Discussion

Our findings differ from previously published reports on members of the MDR superfamily by the electronic-, geometric- and functional characteristics of the metal centers. Zn^{2+} has an [Ar] $3d^{10}$ electron configuration and preferentially adopts a tetrahedral coordination geometry, next to the transient formation of trigonal-bipyramidal complexes during ligand exchange.[21] In contrast, Mg^{2+} is completely devoid of d-electrons ([Ne] $3s^0$), while high-spin Mn^{2+} only contains five half-filled d-orbitals ([Ar] $3d^5$). Both metals generally favor the formation of octahedral complexes, with the additional observation of square-pyramidal complexes in *Sw*HKA. In MDRs and D-xylose isomerase, the dynamic properties of the metal cofactors constitute a key element of their respective catalytic mechanisms, which are required for the reactions to be reversible. In contrast, the movement between $M^{2+}{}_R$ and $M^{2+}{}_{A,S}$ does not serve a mechanistic purpose in *Sw*HKA, but corresponds to a binary switch between two states of distinct properties. While the natural function of the $M^{2+}{}_R$ configuration could not yet be determined, an environment responsive regulation between catalytic activity and enhanced stability is conceptually plausible for this mechanism. Additionally, the requirement for bidentate substrate coordination to transform *Sw*HKA into its catalytically active state constitutes a powerful prerequisite, that increases its substrate specificity. This mechanism of control could find potential applications in the design of artificial metalloenzymes. Here, the bio-compatibility of the (organo-)metallic catalyst is considerably improved by the protein environment, which then allows for its modular incorporation as part of multi-step enzymatic cascade reactions.[22-28] Similar to the regulation of metabolism *in vivo*, control of enzyme activity and selectivity becomes of great importance with increasing complexity of the synthetic cascade reaction. The substrate induced transition from a resting- into a catalytically active state could therefore increase their substrate specificity and constitute an interesting alternative to the design of allosteric regulation in artificial metalloenzymes.[10]

5.4 Conclusion

The structural characterization of *Sw*HKA revealed a movement of its metal cofactor between two coordination sites of distinct properties, which are interconnected by a dynamic equilibrium. The transition from M^{2+}_{R} to $M^{2+}_{A,S}$ is induced by bidentate substrate binding and occurs without a concomitant rearrangement of the protein scaffold. The equilibrium distribution between the individual states is determined by their relative energy difference, which can be modulated both *via* reversible ligand exchange and the use of different metals. This mechanism therefore constitutes an alternative to allosteric regulation for the adjustment of metalloenzyme properties in response to a changed external environment. The conceptual simplicity of a dynamic equilibrium between different metal binding sites suggests that this behavior may be a more widespread phenomenon in nature, which has been largely overlooked until now due to its conditional character and the low occupancy of some of its states in protein crystal structures.

Acknowledgements

Financial support from the ERA-IB scheme (grant ERA-IB-15-110) to L.M. is gratefully acknowledged. Calculations were performed at the peregrine computer cluster at the University of Groningen.

5.5 Supporting information

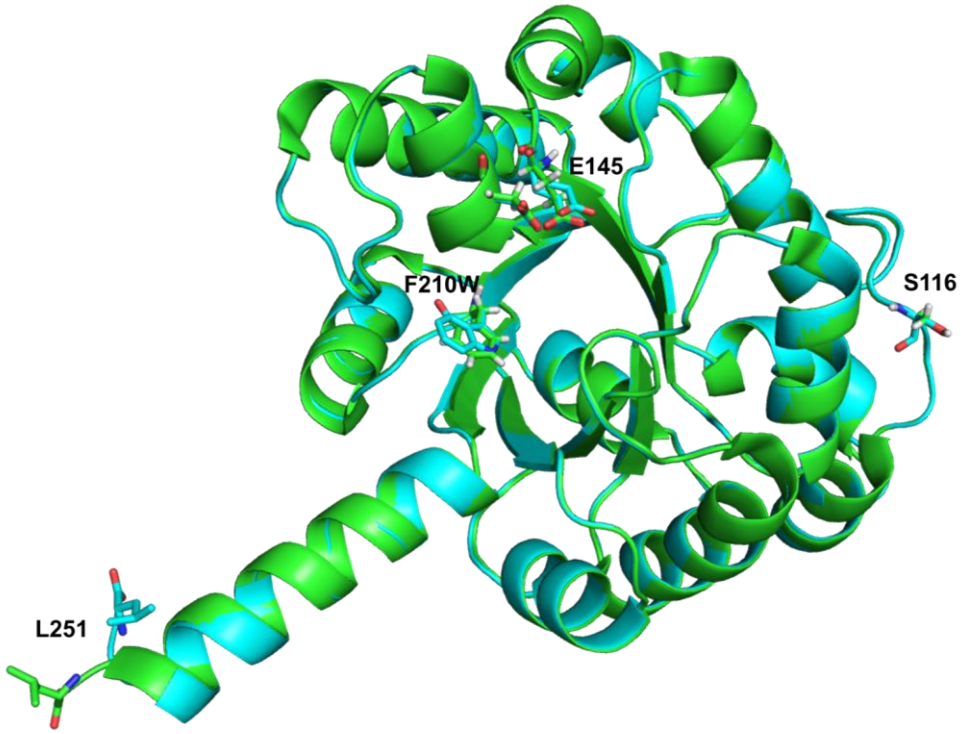


Figure S1: Alignment of WT *Sw*HKA $Mg^{2+}_{A,S}$ (green) and molecule B in F210W *Sw*HKA containing $Mg^{2+}_{A,S}$ (cyan) (rmsd = 0.194 Å). The positions of side chain residues in *Sw*HKA are unaffected by the mutation F210W. The lower occupancy of $Mg^{2+}_{A,S}$ in F210W (0.7) compared to WT *Sw*HKA (1.0) therefore does not seem to have structural reasons, but presumably originates from electronic interactions (such as CH-π) interactions with W210. F210W *Sw*HKA shows residue E145 in the position of M^{2+}_{R} with an occupancy of 0.3, next to a solvent exposed residue L251 as the only differences between both structures. Alignments and figures were made with PyMol.

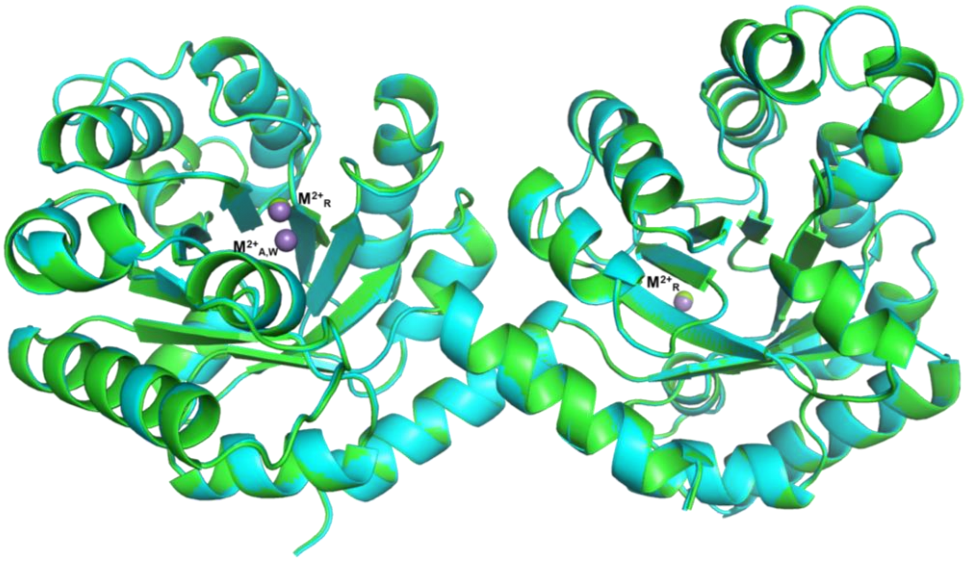


Figure S2: Alignment of WT *Sw*HKA Mg^{2+}_{R} and WT SwHKA $Mn^{2+}_{A,V}/Mn^{2+}_{R}$ (rmsd = 0.164 Å). The use of either $MgCl_2$ or $MnCl_2$ for holoenzyme formation does not alter the overall protein structure. The mixed occupancy of $Mn^{2+}_{A,V}$ and Mn^{2+}_{R} in WT *Sw*HKA (and therefore the relative energy difference between the corresponding metal complexes) is therefore determined by the respective choice of metal, and not due to possible changes in protein structure. Alignments and figures were made with PyMol.

5.6 Methods Summary

Expression and purification. The codon optimized gene was recombinantly expressed in *E.coli* BL21(DE3) cells and the enzyme was subsequently purified by affinity chromatography *via* its *N*-terninal His_6-tag, followed by size exclusion chromatography as previously described.[15]

Crystallography. Protein crystals of *apo*-*Sw*HKA were obtained at 277 K using the vapor diffusion method, from drops resulting from mixing equal volumes of a protein sample at 9.2 mg/ml with a crystallization solution consisting of 0.44 M to 0.65 M of sodium citrate and 0.1 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) at pH 7.05. Crystals took 7 days to appear and grew to their final size within a few days. Crystals of *holo*-*Sw*HKA in complex with hydroxypyruvate were obtained by transferring crystals of *apo*-*Sw*HKA into a

new drop containing 0.7 M of potassium bromide (KBr), 0.1 M HEPES pH 7.05, 20 mM of magnesium chloride ($MgCl_2$) and 150 mM of hydroxypyruvate. The crystals were left in this new condition for a few hours and then were cryo-cooled using as cryoprotectant 20% of glycerol. X-ray diffraction data was collected at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany).[29] The protein crystallized in the cubic space group P 42 3 2, with unit cell parameters of approx. a=116.6 Å, with one molecule in the asymmetric unit and a solvent content of 47.5%.

Data were indexed and integrated with XDS,[30] scaled with AIMLESS,[31] and the space group was determined with POINTLESS[32] from CCP4 programs suite.[33] The structure was solved by the molecular-replacement method with the program MOLREP[34] using the crystal structure of *apo-Sw*HKA as a search model. Model building and refinement of ligand bound *holo-Sw*HKA was performed with COOT[35] and PHENIX.[36] The program PHENIX was used to refine atomic coordinates together with individual isotropic atomic displacement parameters. TLS thermal anisotropic parameterization was also included in the final stages of refinement, with each monomeric subunit divided into three TLS groups, as suggested by TLS Motion Determination.[35,37] A randomly selected 5% of observed reflections were kept aside during refinement for cross-validation. The positions of the magnesium ion and of the hydroxypyruvate became evident after the first cycles of refinement, based on electron density difference maps, and both entities were modelled and refined with full occupancy. Solvent molecules were included in the model after a few rounds of refinement as well as bromide and potassium ions that were located at the surface of the protein.

QM energy calculations. Input structures for the theoretical calculations were generated from the corresponding X-ray structures. An octahedral active form with four water molecules was not observed during crystallography. Therefore, this hypothetical form was created from a substrate bound structure by converting the two coordinating substrate oxygens to water molecules. To evaluate whether hydroxyl ions or water molecules were present as ligands, the H-bonding network inside the active site was predicted with YASARA using the commands "CleanAll" and "OptHydAll".[38,39] There were no indications for hydroxyl ions, and therefore all coordinating water molecules were modeled as neutral. Input structures for the Mn^{2+} complexes were generated from the corresponding Mg^{2+} forms by replacing the metal ion. Calculations were performed with GAUSSIAN09.[40] The APFD and B3LYP-D3[41,42] methods were used separately to compare the results. In all cases, a 6-311++G(2d,p) basis set was used during both energy minimizations and frequency calculations. The solvent was modeled with a polarized continuum model. Because of the dense H-bonding network around the metal ligands inside the active site, water was selected as the solvent. An example of the applied keywords is "#p opt=(CalcFC,tight) freq b3lyp/6-311++g(2d,p) EmpiricalDispersion=GD3 scrf=(iefpcm,solvent=water) Int=UltraFine"(this was used for a Mg^{2+} complex modeled with restricted B3LYP-D3). For the Mn^{2+} complexes, separate calculations were run for the S=1/2 and S=5/2 spin states. None of the final optimized conformations displayed imaginary frequencies. The final reported energies correspond to the sum of electronic and thermal free energies.

Determination of dissociation constants by isothermal calorimetry measurements (ITC). The ITC cell (VP-ITC MicroCal, Malverne) was filled with a stock solution of apo-*Sw*HKA (2 mL, 8.1 mg/mL in 20 mM TEOA, pH 7.5, ≈ 300 µM corresponding to approximately 50x K_d) and titrated against a $MnCl_2$ solution (2.54 mM, 20 mM TEOA, pH 7.5) using identical buffers with the following parameters: number of injections: 28, cell temperature: 25°C, initial delay: 60 seconds. A control experiment was performed to correct for the integrated heat of dilution by titrating the $MnCl_2$ solution into the buffer.

EPR measurements. EPR measurements were carried out using a Bruker EMXplus 9.5 spectrometer and the following conditions: 9.402 GHz microwave frequency, 0.2 mW microwave power, 100 kHz modulation frequency, 10 Gauss modulation amplitude at a temperature of 37 K. The microwave power was optimized by recording a 2D powerplot from 0 to 40 dB using the Xenon software (Bruker), which was analyzed using a labview EPR analysis program written by W.R. Hagen (to be published elsewhere). The low temperature was maintained by boiling liquid helium and the cold helium vapor was passed through a double wall quartz glass tube which was mounted and fitted in the rectangular cavity.[43,44] Samples were prepared in TEOA buffer (5 mM, pH=7.5) to contain 1.2 mM of purified *apo*-enzyme (calculated for the molecular weight of the monomer of 27.4 kDa), Mn^{2+} (1 mM, 0.85 eq.), KPi (2 mM, 1.7 eq.) ketoacids/aldehyde substrate (10 mM, 8.3 eq.) to afford a final volume of 200 µL.

Aldolase activity assay. Purified *apo-Sw*HKA (0.05 mg/mL) was incubated with NADH (0.5 mM), LDH (10 U/mL) in TEOA buffer (20 mM, pH 7.5) and the holoenzyme was formed by the addition of $MnCl_2$ (0.2 mM) and potassium phosphate (5 mM). The reaction was initiated by the addition of oxaloacetate (0.5 mM) to afford a final volume of 1 mL and the change in absorption was followed over time (λ = 340 nm, 25°C, 800 rpm, on a Cary 60 UV-Vis spectrometer (Agilent Technologies), equipped with a TC1 stirring unit (Quantum Northwest)). Specific activities were calculated from the linear slopes and kinetic parameters were calculated by fitting the data with IGOR.

References

[1] Waldron, K.J.; Robinson, N.J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **2009**, *7*, 25.

[2] Dupont, C.L.; Yang, S.; Palenik, B.; Bourne, P.E. Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proceedings of the National Academy of Sciences* **2006**, *103*, 17822-17827.

[3] Andreini, C.; Bertini, I.; Rosato, A. A hint to search for metalloproteins in gene banks. *Bioinformatics* **2004**, *20*, 1373-1380.

[4] Bertini, I.; Cavallaro, G. Metals in the “omics” world: copper homeostasis and cytochrome c oxidase assembly in a new light. *JBIC J. Biol. Inorg. Chem.* **2008**, *13*, 3-14.

[5] Rees, D.C. Great metalloclusters in enzymology. *Annual review of biochemistry* **2002**, *71*, 221-246.
[6] Tainer, J.A.; Roberts, V.A.; Getzoff, E.D. Metal-binding sites in proteins. *Current Opinion in Biotechnology* **1991**, *2*, 582-591.
[7] Goodey, N.M.; Benkovic, S.J. Allosteric regulation and catalysis emerge via a common route. *Nature chemical biology* **2008**, *4*, 474-482.
[8] Laskowski, R.A.; Gerick, F.; Thornton, J.M. The structural basis of allosteric regulation in proteins. *FEBS letters* **2009**, *583*, 1692-1698.
[9] Kern, D.; Zuiderweg, E.R. The role of dynamics in allosteric regulation. *Current opinion in structural biology* **2003**, *13*, 748-757.
[10] Fastrez, J. Engineering allosteric regulation into biological catalysts. *ChemBioChem* **2009**, *10*, 2824-2835.
[11] Kovalevsky, A.Y.; Hanson, L.; Fisher, S.Z.; Mustyakimov, M.; Mason, S.A.; Forsyth, V.T.; Blakeley, M.P.; Keen, D.A.; Wagner, T.; Carrell, H. Metal ion roles and the movement of hydrogen during reaction catalyzed by D-xylose isomerase: a joint x-ray and neutron diffraction study. *Structure* **2010**, *18*, 688-699.
[12] Baker, P.J.; Britton, K.L.; Fisher, M.; Esclapez, J.; Pire, C.; Bonete, M.J.; Ferrer, J.; Rice, D.W. Active site dynamics in the zinc-dependent medium chain alcohol dehydrogenase superfamily. *Proceedings of the National Academy of Sciences* **2009**, *106*, 779-784.
[13] Maret, W.; Li, Y. Coordination dynamics of zinc in proteins. *Chemical reviews* **2009**, *109*, 4682-4707.
[14] De Berardinis, V.; Guérard-Hélaine, C.; Darii, E.; Bastard, K.; Hélaine, V.; Mariage, A.; Petit, J.-L.; Poupard, N.; Sánchez-Moreno, I.; Stam, M. Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chemistry* **2017**, *19*, 519-526.
[15] Marsden, S.R.; Mestrom, L.; Bento, I.; Hagedoorn, P.L.; McMillan, D.G.; Hanefeld, U. CH-π Interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase. *Advanced Synthesis & Catalysis* **2019**, *361*, 2649-2658.
[16] Fang, J.; Hait, D.; Head-Gordon, M.; Chang, M. Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angew. Chem.* **2019**, *131*, 11967-11971.
[17] Laurent, V.; Uzel, A.; Hélaine, V.; Nauton, L.; Traikia, M.; Gefflaut, T.; Salanoubat, M.; de Berardinis, V.; Lemaire, M.; Guérard-Hélaine, C. Exploration of Aldol Reactions Catalyzed by Stereoselective Pyruvate Aldolases with 2-Oxobutyric Acid as Nucleophile. *Adv. Synth. Catal.* **2019**, *361*, 2713-2717.
[18] Siegbahn, P.E.; Borowski, T. Modeling enzymatic reactions involving transition metals. *Accounts of chemical research* **2006**, *39*, 729-738.
[19] Chen, S.-L.; Fang, W.-H.; Himo, F. Technical aspects of quantum chemical modeling of enzymatic reactions: the case of phosphotriesterase. *Theoretical Chemistry Accounts* **2008**, *120*, 515-522.

[20] Vertregt, F.; Torrelo, G.; Trunk, S.; Wiltsche, H.; Hagen, W.R.; Hanefeld, U.; Steiner, K. EPR study of substrate binding to Mn (II) in hydroxynitrile lyase from Granulicella tundricola. *ACS Catalysis* **2016**, *6*, 5081-5085.
[21] Laitaoja, M.; Valjakka, J.; Jänis, J. Zinc coordination spheres in protein structures. *Inorganic chemistry* **2013**, *52*, 10983-10991.
[22] Liu, Z.; Lebrun, V.; Kitanosono, T.; Mallin, H.; Köhler, V.; Häussinger, D.; Hilvert, D.; Kobayashi, S.; Ward, T.R. Upregulation of an artificial zymogen by proteolysis. *Angew. Chem. Int. Ed.* **2016**, *55*, 11587-11590.
[23] Jing, Q.; Okrasa, K.; Kazlauskas, R.J. Stereoselective Hydrogenation of Olefins Using Rhodium-Substituted Carbonic Anhydrase—A New Reductase. *Chemistry–A European Journal* **2009**, *15*, 1370-1376.
[24] Jing, Q.; Kazlauskas, R.J. Regioselective Hydroformylation of Styrene Using Rhodium-Substituted Carbonic Anhydrase. *ChemCatChem* **2010**, *2*, 953-957.
[25] Fujieda, N.; Nakano, T.; Taniguchi, Y.; Ichihashi, H.; Sugimoto, H.; Morimoto, Y.; Nishikawa, Y.; Kurisu, G.; Itoh, S. A well-defined osmium–cupin complex: hyperstable artificial osmium peroxygenase. *J. Am. Chem. Soc.* **2017**, *139*, 5149-5155.
[26] Zhao, J.; Kajetanowicz, A.; Ward, T.R. Carbonic anhydrase II as host protein for the creation of a biocompatible artificial metathesase. *Organic & biomolecular chemistry* **2015**, *13*, 5652-5655.
[27] Heinisch, T.; Ward, T.R. Latest developments in metalloenzyme design and repurposing. *Eur. J. Inorg. Chem.* **2015**, *2015*, 3406-3418.
[28] Fessner, W.-D. Systems Biocatalysis: Development and engineering of cell-free “artificial metabolisms” for preparative multi-enzymatic synthesis. *New biotechnology* **2015**, *32*, 658-664.
[29] Cianci, M.; Bourenkov, G.; Pompidor, G.; Karpics, I.; Kallio, J.; Bento, I.; Roessle, M.; Cipriani, F.; Fiedler, S.; Schneider, T.R. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high-and low-energy phasing with variable beam focusing. *Journal of synchrotron radiation* **2017**, *24*, 323-332.
[30] Kabsch, W. Xds. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 125-132.
[31] Evans, P.R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 282-292.
[32] Evans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography* **2013**, *69*, 1204-1214.
[33] Collaborative, C.P. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **1994**, *50*, 760.
[34] Vagin, A.; Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of applied crystallography* **1997**, *30*, 1022-1025.

[35] Painter, J.; Merritt, E.A. A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. *Acta Crystallographica Section D: Biological Crystallography* **2005**, *61*, 465-471.

[36] 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. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 213-221.

[37] Winn, M.; Isupov, M.; Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica Section D: Biological Crystallography* **2001**, *57*, 122-133.

[38] Krieger, E.; Dunbrack, R.L.; Hooft, R.W.; Krieger, B. Assignment of protonation states in proteins and ligands: Combining pK a prediction with hydrogen bonding network optimization. In *Computational Drug Discovery and Design*, Springer: 2012; pp. 405-421.

[39] Krieger, E.; Vriend, G. YASARA View—molecular graphics for all devices—from smartphones to workstations. *Bioinformatics* **2014**, *30*, 2981-2982.

[40] Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G., et al. Gaussian 09; Gaussian, Inc. *Wallingford, CT* **2009**, *32*, 5648-5652.

[41] Austin, A.; Petersson, G.A.; Frisch, M.J.; Dobek, F.J.; Scalmani, G.; Throssell, K. A density functional with spherical atom dispersion terms. *Journal of chemical theory and computation* **2012**, *8*, 4989-5007.

[42] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics* **2010**, *132*, 154104.

[43] Salmeen, I.; Palmer, G. Electron Paramagnetic Resonance of Beef-Heart Ferricytochrome c. *Journal of Chemical Physics* **1968**, *48*, 2049-2052.

[44] Lundin, A.; Aasa, R. A simple device to maintain temperatures in the range 4.2–100 K for EPR measurements. *Journal of Magnetic Resonance* **1972**, *8*, 70-73.

6

Investigation of an Aldolase/Transaminase Cascade Reaction for the Synthesis of C-3 Modified Amino Acids

Aldolase/transaminase cascade reactions allow for the selective, one-pot synthesis of non-natural amino acids from simple building blocks. The incorporation of recently discovered promiscuous pyruvate aldolases could potentially give modular access to C-3 modified amino acids. In this study, 102 wild-type transaminases were screened for the conversion of D-ketogluconic acid as a C-3 hydroxylated model compound. This revealed two main challenges: cyclic pyranose and furanose forms of the carbohydrate-like substrate render the reaction thermodynamically unfavourable, while a low catalytic efficiency of the screened wild-type enzymes makes it challenging from a kinetic standpoint. These issues could be addressed by directed evolution of e.g. the phosphoserine aminotransferase SerC to increase the catalytic efficiency. Alternatively, a different choice of aldehyde substrates could prevent the formation of cyclic intermediates and thereby establish a more favourable thermodynamic equilibrium.

This work was conducted under the local supervision of Prof. Helen Hailes and Prof. John Ward at the Departments of Chemistry and Biochemical Engineering at UCL from May to July 2019. Experiments were conducted in collaboration with Dr. Leona Leipold, Dr. Fabiana Subrizi and Dr. Max Cardenás-Fernández.

6.1 Introduction

New biocatalytic conversions are traditionally discovered by screening a large number of enzyme preparations for the desired activity.[1] This approach has been further refined using *in silico* methods, such as database mining and metagenomics,[2-4] next to directed evolution strategies[5-8] and *de novo* computational design.[9,10] Taken together, these efforts have established a considerable biocatalytic toolbox.[11] However, many chemical conversions still lack a biocatalytic analogue, since the synthetic scope of enzymes remains limited by the chemical properties of the canonical amino acids.[12]

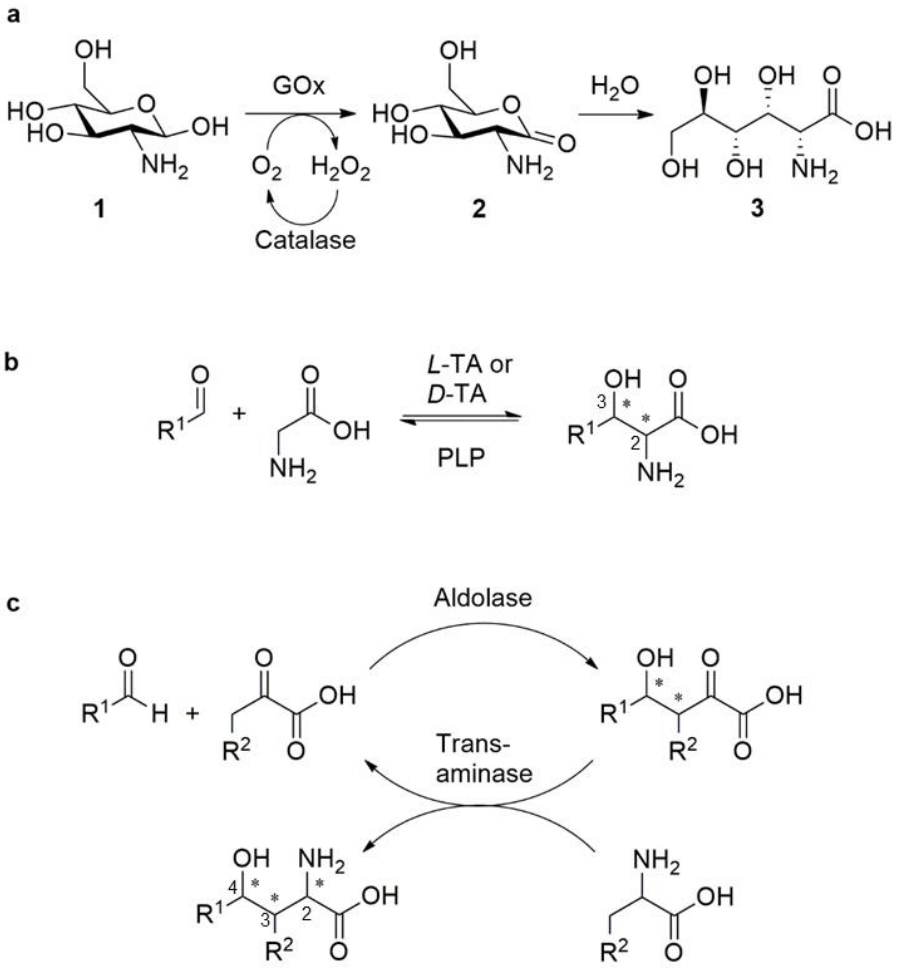
The incorporation of non-natural amino acids with unusual chemical properties further expands applications of the biocatalytic toolbox with new-to-nature reactions of synthetic relevance.[12] Due to the large redundancy of natural three letter codons (64 possible codons for 20 amino acids and 3 stop codons), variants of redundant aminoacyl-tRNA synthetases can be used for the introduction of non-natural amino acids during ribosomal gene expression.[13] The design of four[14,15] and five-base codons has further expanded possibilities to the simultaneous incorporation of non-natural amino acids.[16] The development of efficient synthetic routes towards diversely functionalised, non-natural amino acids is therefore an important step towards many potentially new applications within biotechnology.

Hydroxylated amino acids are involved in the formation of hydrogen bond networks, charge delocalization,[17] complexation of metal cofactors,[18] and function as a nucleophile in serine hydrolases.[19,20] The introduction of polyhydroxylated amino acids could therefore enable some interesting new design strategies for rational mutagenesis. In this chapter, an aldolase/transaminase cascade reaction is investigated for the selective synthesis of C-3 modified polyhydroxy amino acids.

Biocatalytic synthesis strategies for polyhydroxy amino acids

2-Amino sugars are naturally abundant *via* the transamination of ketose phosphates or ketose nucleotides.[21] Subsequent oxidation of the anomeric carbon affords the corresponding 2-amino acids (Scheme 1a), in which the product stereochemistry is predetermined by the configuration of the substrate.[22] While glucose oxidase (GOx) catalyses the oxidation of D-glucosamine **1** to D-glucosaminic acid **3**,[23,24] no other 2-amino sugars are accepted by GOx. Threonine aldolases (TAs) utilise pyridoxal-5'-phosphate (PLP) and a divalent metal as cofactors to afford β-hydroxy-α-amino acids from glycine and aldehydes, such as glycolaldehyde,[25] benzaldehyde derivatives[26] and aliphatic aldehydes (Scheme 1b).[27] The Schiff base intermediate between glycine and the PLP cofactor affords a good chromophore, which can be used for the spectrophotometric screening of donor substrates other than glycine.[28] TAs display a high stereoselectivity at the α-carbon, and stereocomplementary enzymes have been identified to afford both D- and L-configured amino acids. While TAs are only moderately selective at the β-carbon, both (*R*) and (*S*) configured products can be isolated under either thermodynamically or kinetically controlled conditions.[29] TAs thus allow for enzymatic stereocontrol at C-2 and C-3.

Aldolase/transaminase cascade reactions are a promising approach for the control of stereochemistry at three positions (C-2 by the transaminase, C-3 and C-4 by the aldolase, Scheme 1c) through the modular combination of stereocomplementary enzymes. In its design as a cascade, the ketoacid substrate can be generated *in situ* from the corresponding amino acid, which addresses a major cost driver in the case of hydroxypyruvate.[30,31] The feasibility of pyruvate aldolase/transaminase based cascade systems has previously been demonstrated with alanine (R^2 = H) and formaldehyde as substrates, to afford 2-amino-4-hydroxybutanoic acid.[32] The use of promiscuous pyruvate aldolases[33] could allow for the stereoselective introduction of structural diversity at C-3 (R^2 = OH, Me, F) in a similar cascade system.

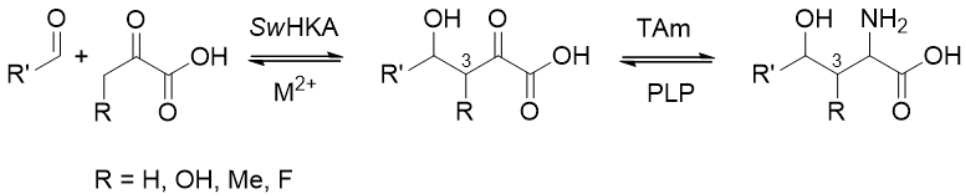


Scheme 1: (a) Oxidation of D-glucosamine to D-glucosaminic acid by glucose oxidase. The stereochemistry is determined by the configuration of the substrate. **(b)** Threonine aldolase catalysed reactions allow for control of stereochemistry at C-2 and C-3. **(c)** Two-step aldolase/transaminase based cascades theoretically allow for control of stereochemistry at C-2,C-3,C-4 and structural diversification at C-3 (R^2 = H, OH, Me, F).

The class II pyruvate aldolase from *Sphingomonas Wittichii RW1* (*Sw*HKA, Uniprot ID A5VH82) accepts hydroxypyruvate (HPA) as a donor substrate to afford

(*3S*,*4S*) configured dihydroxy ketoacids as aldol products.[33] Further investigations into its donor substrate scope revealed the efficient conversion of 2-oxobutyrate[34] and fluoropyruvate into the corresponding C-3 methylated or fluorinated aldol products.[35] On the acceptor side of the aldol reaction, *Sw*HKA converts aliphatic-, aromatic- and hydroxylated aldehydes[33,35] next to several ketones.[36] A considerable library of structurally diverse ketoacids can therefore be synthesised by *Sw*HKA, which makes it an interesting enzyme candidate for applications in an aldolase/transaminase cascade system.

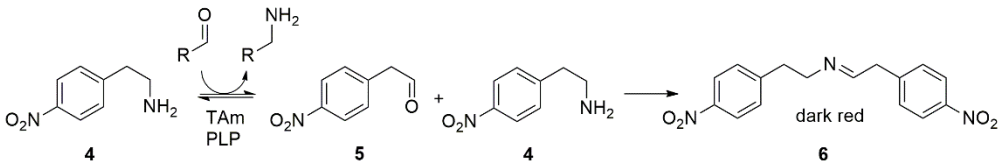
Scheme 2 shows a simplified version of the envisioned aldolase/transaminase cascade reaction, in which the two-step conversion occurs sequentially. Since the corresponding aldol reactions are well established on a preparative scale,[33,35,36] we chose to focus our efforts on the screening of transaminases for the second step, converting a hydroxy ketoacid into the corresponding amino acid. D-ketogluconic acid is formed as the aldol product of hydroxypyruvate with D-glyceraldehyde under kinetically controlled conditions in the case of *Bp*HKA (Uniprot ID B2T1L6) or under thermodynamically controlled conditions with *Sw*HKA.[33] Due to its commercial availability, the hemicalcium salt of D-ketogluconic acid was chosen as model compound for this study, with D-glucosaminic acid as a commercial product standard. In the following work, more than one hundred transaminases were screened for the conversion of D-ketogluconic acid to D-glucosaminic acid with different amine donors.



Scheme 2: Sequential aldolase/transaminase cascade reaction using promiscuous pyruvate aldolases for structural diversification at C-3.

6.2 Results & Discussion

A set of nine previously characterised transaminases from the UCL culture collection including both (*R*) and (*S*) selective variants (Table 1), were heterologously expressed in *Escherichia coli* BL21(DE3). Cell-free lysates were prepared in the range of 10-20 mg/mL of soluble, total protein and transaminase activity was evaluated by a colorimetric assay using sodium pyruvate and 4-nitrophenylethylamine **4** as substrates (Scheme 3).[2] The transaminase from *Mycobacterium vanbaalenii* (MV, UniProt ID A1TDP1) was only poorly expressed (Figure 7), but qualitatively displayed good transaminase activity in the colorimetric assay (Figure 8). The two thermostable transaminases from *Deinococcus geothermalis* (pQR1757, UniProt ID Q1IZC2) and *Geobacillus stearothermophilus* (pQR1756, UniProt ID Q59228) required an extended incubation for 10 minutes in boiling SDS sample buffer for complete denaturation of the homodimer prior to SDS-PAGE analysis (Figure 7). In agreement with previous reports, the two thermostable transaminases did not accept 4-nitrophenylethylamine as amine donor,[31] while all seven other enzymes showed moderate to good transaminase activity with **4** as substrate. However, no conversion of D-ketogluconate was observed by the colorimetric assay during a first screening using an enzyme loading of 0.4-0.65 mg/mL total protein (Figure 1).

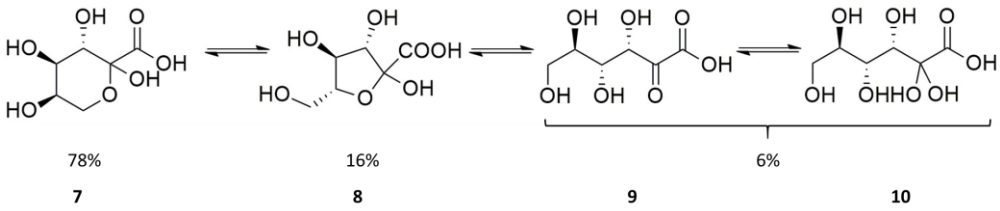


Scheme 3: Colorimetric assay for the screening of transaminase activity using 4-nitrophenylethylamine **4** as amine donor. Upon reaction, the formed 4-nitrophenylacetaldehyde **5** spontaneously reacts with residual starting amine **4**, resulting in a red colour change as an indicator of transaminase activity.[2]

Equilibrium of D-ketogluconate isomers in aqueous solution

In aqueous solution, D-ketogluconic acid **9** is in equilibrium with the ketals **7** and **8** next to its hydrated form **10**. This leads to a lower concentration of its linear ketoacid form, which is the active form in transaminase reactions. The relative concentrations of each isomer were therefore determined under standard reaction conditions by ^{1}H-NMR spectroscopy. The pyranose (**7**, 78%) and furanose forms (**8**, 16%) constitute the major species, and the two forms were discriminated *via* their corresponding HMBC cross peaks ($^{3}J_{H6C2}$ is characteristic for pyranoses). The ^{1}H- and ^{13}C-NMR signals were fully assigned for the pyranose and furanose forms using ^{1}H^{1}H-COSY, HSQC and HMBC NMR spectrums.

In contrast, the linear ketone **9** and hydrate **10** were only observed in low quantities (6% of total substrate). Due to their low concentration and largely similar chemical shifts, the signals could not be fully discriminated and assigned under the conditions used (700 MHz for ^{1}H). Hydroxy ketoacids are highly electrophilic in nature and are predominantly observed in their hydrated form; the amount of free ketone was therefore estimated to be in the range of 1% of total substrate. While dynamic equilibriums in principle allow for the complete conversion of a substrate *via* its minor isomer, problems may arise from an insufficient affinity of the enzyme towards the substrate at low concentrations.



Scheme 4: Equilibrium distribution of D-ketogluconic acid isomers and hydrated form in aqueous solution, as determined by ^{1}H-NMR spectroscopy.

Modulation of the water activity with DMSO as a co-solvent previously indicated a beneficial shift of the equilibrium distribution towards the linear isomer for carbohydrates.[31] However, no observable change in concentration of the linear isomers was observed by NMR spectroscopy in the presence of 10% DMSO-D_6, while 20% DMSO-D_6 negatively affected the quality of the spectrum, no longer allowing for a reliable integration (Figure 22). Nonetheless, the colorimetric assay was repeated with 10-40% (v/v) DMSO as co-solvent and six-fold increased enzyme loadings (2.8-3.3 mg/mL total protein). All previously active enzymes retained their activity in the presence of 10% DMSO using pyruvate as a positive control. While an apparent darker colour was indicative of conversion, the concomitant increase in background response from the high concentration of cell lysate did not allow for the definite confirmation of transaminase activity towards D-ketogluconate.

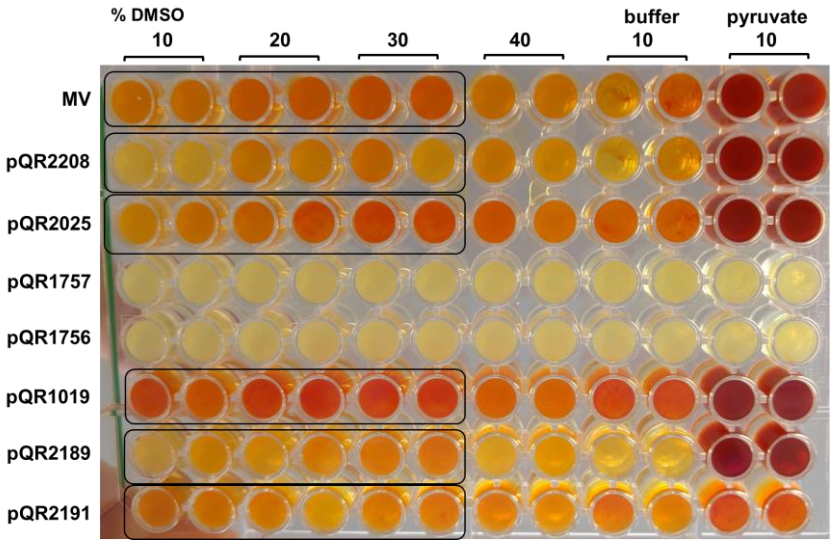
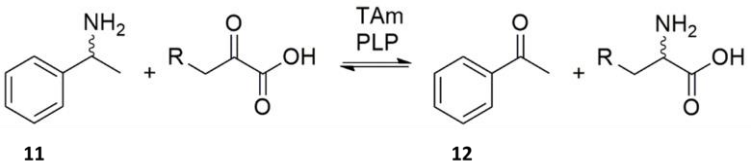


Figure 1: Representative result of the colorimetric assay, screening for conversion of D-ketogluconate. Conditions: *holo*-transaminase (0.4-3.3 mg/mL total protein, 0.5 mM PLP, 10 mM ketoacid substrate, 25 mM 4-nitrophenyl-2-ethylamine·HCl, 30 °C, 500 rpm, 24 h). Transaminase activity is indicated by a red colour change. Enzyme IDs are shown in table 1.

We therefore switched to a quantitative assay using the appropriate enantiomer of methylbenzylamine **11** (MBA) as amine donor. Here, conversions were followed *via* the stoichiometric formation of acetophenone **12** (Scheme 5).



Scheme 5: Methylbenzylamine (MBA) assay using either (*R*)-MBA or (*S*)-MBA as amine donor to match the transaminase's stereopreference. Conversion was determined indirectly *via* the formation of acetophenone at 254 nm.

Calcium inhibition of transaminase activity

While >90% conversion was observed with pyruvate as an acceptor, no conversion was observed for D-ketogluconate. Notably, lower levels of background activity were observed in the presence of D-ketogluconate hemicalcium, indicating substrate inhibition. TAm pQR2189 showed the lowest background activity and was therefore investigated in inhibition studies. Transamination reactions of pyruvate were supplemented with $CaCl_2$ or D-ketogluconate hemicalcium. Significantly lower rates were observed in the presence of Ca^{2+} (Figure 2). Notably, no difference was observed when chloride or ketogluconate were used as counterions. Qualitative removal of Ca^{2+} from the reaction mixture was achieved by precipitation as insoluble calcium phosphate. This restored the initial, uninhibited progress curve. In summary, these findings suggest the reversible inhibition of transaminase pQR2189 by Ca^{2+}, while the presence of chloride or D-ketogluconate as counterions did not affect the enzyme. Since D-ketogluconate is only commercially available as the hemicalcium salt, precipitation of calcium as insoluble calcium phosphate was used as a rapid ion exchange method for the preparation of stock solutions from here on.

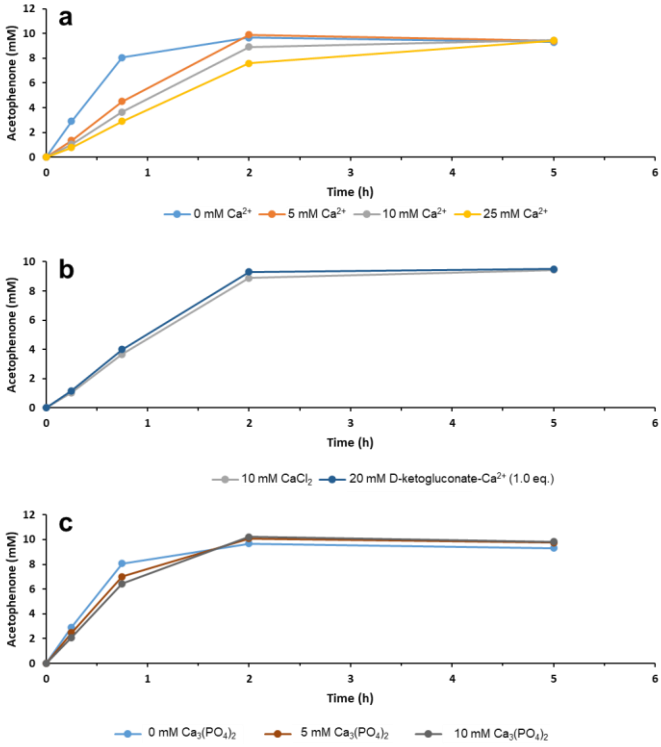
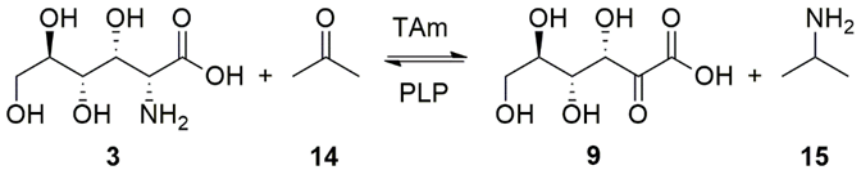


Figure 2: (a) Inhibitory effect of Ca^{2+} on the transamination of pyruvate with methylbenzylamine, catalysed by pQR2189 (n=1). **(b)** Inhibition by equimolar amounts of Ca^{2+} using either $CaCl_2$, or D-ketogluconate hemicalcium. The observed inhibition is independent of the nature of the counter anion (n=1). **(c)** Precipitation of Ca^{2+} as insoluble calcium phosphate prior to the addition of substrates restored uninhibited conditions (n=1). Conditions: 0.05 mg/mL pQR2189, 0.5 mM PLP, 10 mM sodium pyruvate, 25 mM (*S*)-MBA , 100 mM HEPES, pH 7.5, 30 °C, 500 rpm.

Elevated concentrations of phosphate buffer are known to deprive enzymes of their metal cofactor due to the low solubility of divalent metal phosphates.[18,37] However, a similar effect of Ca^{2+} on the solubility of the pyridoxal-5'-phosphate cofactor could not be confirmed during solubility studies, where PLP remained fully soluble in the presence of up to 90 mM $CaCl_2$.

Transamination of D-glucosaminic acid

Unlike D-ketogluconic acid, D-glucosaminic acid does not form cyclic isomers in aqueous solution. The target reaction was therefore also investigated in the reverse direction using D-glucosaminic acid and acetone as substrates. To allow for its analysis by HPLC, the amino groups in D-glucosaminic acid and isopropylamine (IPA) were derivatised with fluorenylmethyloxycarbonyl chloride (Fmoc chloride). This increased the sensitivity of detection and allowed for the efficient separation of substrates and products on a C_{18} column by HPLC. Conversion was quantified using a calibration line of IPA-Fmoc as an external standard.



Scheme 6: Transaminase catalysed conversion of D-glucosaminic acid **3** with acetone **14** to afford D-ketogluconic acid **9** and isopropylamine **15**.

First, the tolerance of the transaminases used towards elevated concentrations of acetone was investigated by the colorimetric assay. Reaction mixtures containing sodium pyruvate and 4-nitrophenylethylamine were supplemented with 5-40% (v/v) acetone. Transaminases pQR1005, pQR2191 were inactivated by the presence of 5% acetone (500 mM), while the other enzymes showed moderate to good residual activity at up to 30% (v/v) acetone (Figure 3). Negative controls in the absence of enzyme showed no interference by the presence of acetone with the assay.

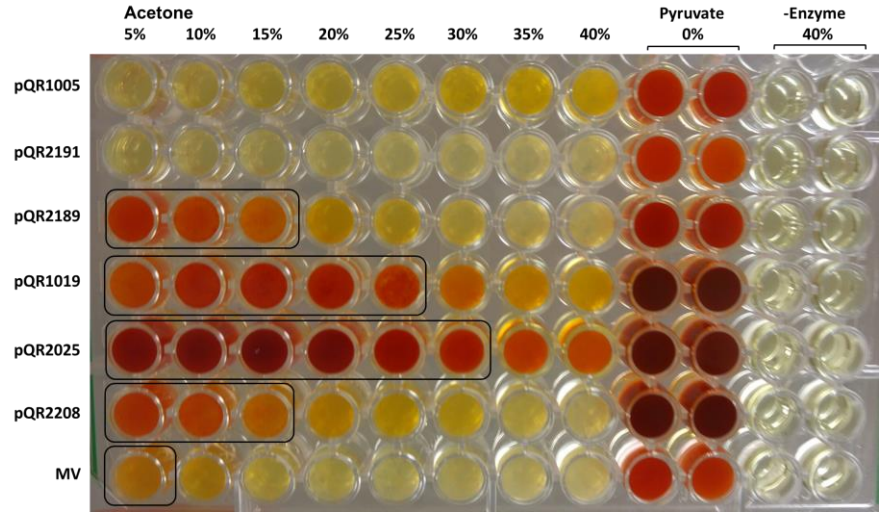


Figure 3: Representative data for the residual activity of selected transaminases in the presence of 5 - 40% (v/v) acetone. Conditions: *holo*-transaminase (0.4-3.3 mg/mL total protein, 0.5 mM PLP, 10 mM sodium pyruvate, 25 mM 4-nitrophenyl-2-ethylamine·HCl, 100 mM HEPES, pH 7.5, 30 °C, 500 rpm, 24 h, n = 1). Enzyme identities are shown in table 1.

The conversion of D-glucosaminic acid was investigated using 25 equivalents of acetone (250 mM). Under these conditions, conversion of D-glucosaminic acid was indicated for all seven transaminases by the formation of low levels of isopropylamine (detected as IPA-Fmoc). Negative controls with empty vector lysates, or in the absence of substrates did not result in the detection of IPA. Notably, comparable levels of conversion were observed for both the known (*2R*) and (*2S*) selective transaminases, thus promising access to both D- and L-configured amino acids (Figure 4). Variation of the enzyme concentration using the *R*-selective transaminase from *Mycobacterium vanbaalenii* (*MV*) showed a linear relationship between conversion and enzyme loading (Figure 5). Cyclisation of D-ketogluconic acid **9** to cyclic ketals **7** and **8** next to the used excess of acetone renders this reaction thermodynamically favourable. The conversion of D-glucosaminic acid must therefore be kinetically limited for all tested enzymes.

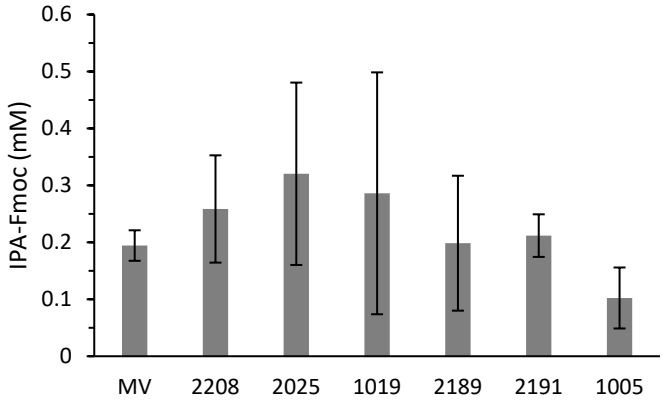


Figure 4: Transamination of D-glucosaminic acid with acetone. Conversions were determined *via* the formation of isopropylamine by HPLC upon derivatisation with Fmoc-chloride. Conditions: *holo*-transaminase (2-3.3 mg/mL total protein, 0.5 mM PLP, 25 mM D-glucosaminic acid, 250 mM acetone, 100 mM HEPES, pH 7.5, 30 °C, 500 rpm, 24 h, n = 1). Enzyme identities are shown in table 1.

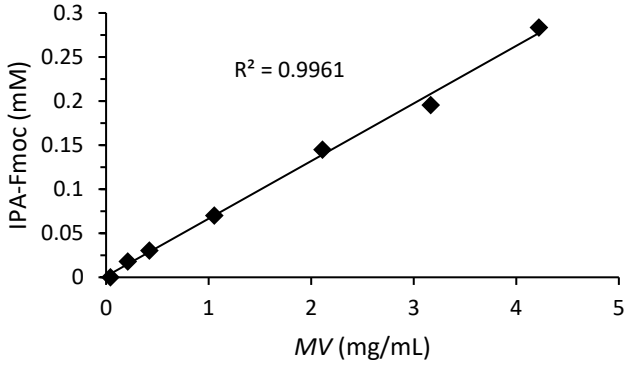


Figure 5: Conversion of D-glucosaminic acid linearly correlated with the enzyme loading of *MV* transaminase. While thermodynamically favourable, the reaction therefore seems to be kinetically limited by the low catalytic efficiency of the tested transaminases towards this non-natural substrate.

Second screening round

Since all enzymes displayed insufficient kinetic properties for the transamination of D-ketogluconate and D-glucosaminic acid, a second screening of 94 additional transaminases was conducted (pQR2501-pQR2594, enzyme ID's are to be published elsewhere by Dr. Max Cárdenas-Fernández).

To facilitate a higher throughput, the enzymes were expressed in 96 deep-well plates. Transaminase activity towards the structural analogues DL-serine/hydroxypyruvate was first confirmed using a previously developed colorimetric assay based on the tetrazolium salt WST-1.[38] Out of the 96 screened enzyme preparations, 81 transaminases displayed good activity towards DL-serine with sodium pyruvate as acceptor substrate. These 81 enzymes were subsequently screened for the conversion of both D-ketogluconic acid and D-glucosaminic acid, using either DL-serine or hydroxypyruvate, respectively. However, no conversion was observed in all cases with D-ketogluconate as substrate (Figure 6b).

Again, only trace amounts of serine (≈ 0.1% conversion) were observed, suggesting the conversion of D-glucosaminic acid in the thermodynamically more favourable reverse direction. The design of the screening workflow in 96-well format for the purpose of higher throughput entailed a decreased enzyme loading. This explains the lower levels of conversion that were observed during the second screening round for D-glucosaminic acid. In summary, all 81 active enzymes of the second screening round similarly displayed an insufficient catalytic efficiency towards these non-natural substrates (Figure 6c).

This challenge could be addressed by directed evolution to increase both affinity (K_M) and activity (k_{cat}) towards the 3,4-dihydroxy-2-oxoacid structural motif starting from a suitable enzyme scaffold.

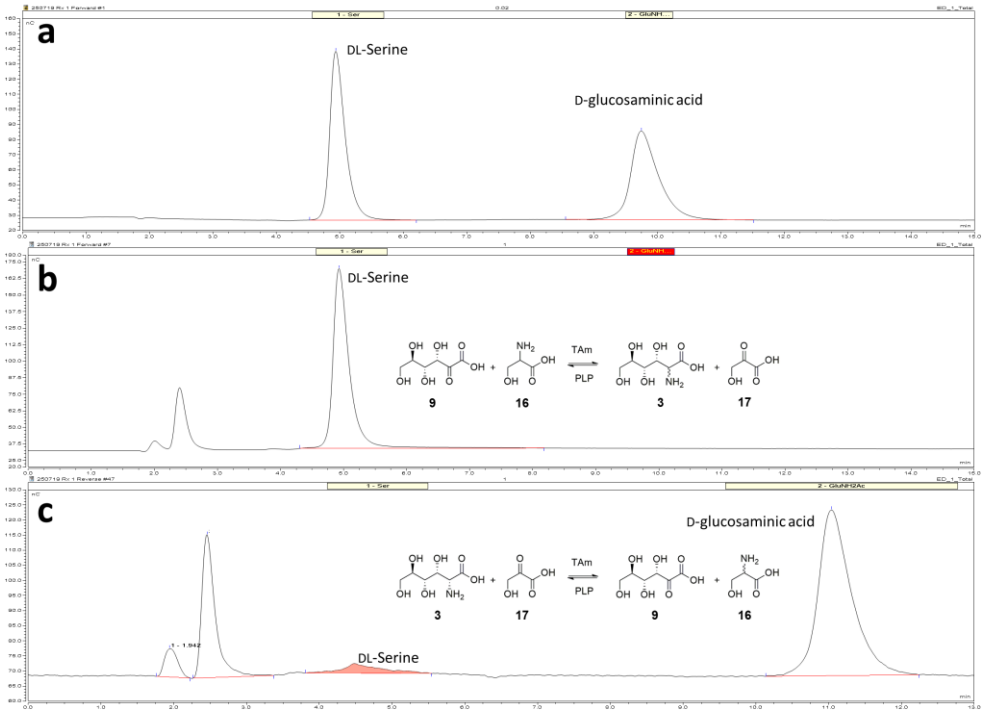


Figure 6: (a) Separation of DL-serine and D-glucosaminic acid by ICS. **(b)** Representative reaction of D-ketogluconate with DL-serine. No conversion was observed. Conditions: 20 µL lysate, 20 mM DL-serine, 0.1 mM PLP, 10 mM D-ketogluconate, 100 mM NaPi, pH 7.0, 50 °C, 18 h. **(c)** Representative reaction of D-glucosaminic acid with hydroxypyruvate. Traces of serine were observed (**≈** 0.001 mg/mL **/** 0.1% conversion), indicating conversion. The retention for the peak corresponding to D-glucosaminic acid time shifted from 9.75 min. to 11.0 min. over the course of the analysis. Conditions: 20 µL lysate, 10 mM hydroxypyruvate, 0.1 mM PLP, 10 mM D-glucosaminic acid, 100 mM NaPi, pH 7.0, 50 °C, 18 h.

For instance, the phosphoserine aminotransferase SerC (Uniprot ID P23721) utilises L-4-phosphohydroxythreonine as its natural substrate.[25] While its strict requirement for phosphorylated substrates excluded it from this study,[39] the rational redesign of its active sites towards non-phosphorylated substrates would make SerC an interesting candidate for directed evolution studies.[40-42]

6.3 Conclusions

In this project, a total of 102 transaminases were heterologously expressed in *E.coli*, of which 87 enzymes displayed good transaminase activity. The conversion of D-glucosaminic acid is thermodynamically favourable due to the spontaneous formation of cyclic isomers of the product D-ketogluconic acid. Yet, only low levels of conversion were observed. This suggests a poor catalytic efficiency of all screened transaminases towards the non-natural substrate D-glucosaminic acid, and in extension likewise for the conversion of D-ketogluconic acid. An impaired access of the substrate into the active site is further supported by a complete lack of competitive substrate inhibition by D-ketogluconate in the case of pQR2189.

In general, the use of aliphatic or aromatic aldehydes instead of hydroxyaldehydes as aldol substrates would circumvent the formation of cyclic pyranose or furanose isomers, and overall render the envisioned cascade reaction thermodynamically more favourable. For the synthesis of polyhydroxylated amino acids, enzyme engineering strategies could be applied to suitable enzyme scaffolds, such as SerC, to improve the kinetic properties for the conversion of non-phosphorylated 3,4-dihydroxy-2-oxoacids.

In summary, the characteristic 3,4-dihydroxy structural motif of *Sw*HKA catalysed aldol reactions with hydroxypyruvate was demonstrated to be accepted by both (*R*)- and (*S*)-selective wild-type transaminases, albeit with poor catalytic efficiency. This confirms the initial hypothesis of control over the stereochemistry at three positions (C-2 to C-4) in aldolase/transaminase cascade reactions. However, substantial optimisation of the catalytic system is required to render the overall cascade economically feasible.

Acknowledgements

This work was conducted under the local supervision of Prof. Helen Hailes and Prof. John Ward at the Departments of Chemistry and Biochemical Engineering at University College London, and the project was further remotely supervised by Prof. Ulf Hanefeld and Dr. Duncan McMillan at Delft University of Technology.

I would like to thank Dr. Leona Leipold, Dr. Fabiana Subrizi and Dr. Max Cárdenas-Fernández for their hands-on help in the laboratories and for valuable discussions.

6.4 Materials & Methods

Chemicals: all chemicals were used without further purification. 2-Keto-D-gluconic acid hemicalcium salt (≥99%, Sigma Aldrich), glucosaminic acid (≥98%, Sigma Aldrich), sodium pyruvate (98%, Sigma Aldrich), D-serine (99% Alfa Aesar), L-serine (≥99.5%, Sigma Aldrich), pyridoxal 5'-phosphate (≥98%, Sigma Aldrich), 4-nitrophenylethylamine hydrochloride (>98%, Alfa Aesar), (*R*/*S*) α-methylbenzylamine (98%, Alfa Aesar), isopropylamine (>99% Alfa Aesar), Fmoc chloride (≥98%, Sigma Aldrich Novabiochem), acetone (Fluka, >99.8%).

Instruments: NMR spectra were recorded on a Bruker ascend 700 (operating at 700 MHz for ^{1}H and 176 MHz for ^{13}C) at room temperature using the residual solvent peak as reference. Chemical shifts are reported in ppm relative to tetramethylsilane and coupling constants (*J*) were measured in Hertz. HPLC analyses were performed on an Agilent 1260 Infinity instrument equipped with an Ace 5 C18 (150 x 4.6 mm, ACE) column. Ion chromatography was conducted on a Dionex ICS-5000^{+} system (Thermo Scientific) equipped with a Dionex Aminopac PA1 anion exchange column (4 x 250 mm, fitted with a Dionex Aminopac PA1 guard column 4 x 50 mm) with electrochemical detection and a KOH 500 cartridge for eluent generation. Mass spectra were measured on a Waters LCT Premier XE ESI Q-TOF mass spectrometer. Precast SDS-PAGE gels (Novex WedgeWell 4-20% Tris-Glycine, Invitrogen) were run in a XCell SureLock electrophoresis cell (Life Technologies) with InstantBlue for protein staining (Expedeon).

Enzyme expression: pre-cultures of *E.coli* BL21(DE3) containing the target genes in the pET28 or pET29 plasmid were grown overnight (5 mL TB medium, 50 mL falcon tube, 0.5 µg/mL kanamycin, 37 °C, 140 rpm) and used as inoculum (1 mL culture in 100 mL TB medium, 250 mL baffled Erlenmeyer flask, 0.5 µg/mL kanamycin, 37 °C, 140 rpm). Expression was induced with IPTG (1 mM) at OD_{600} = 0.6-0.8 and carried out overnight (25 °C, 140 rpm). Cells were harvested by centrifugation (10 min. 5000 rpm, 4 °C) and the pellet

was washed with buffer (15 mL, 100 mM HEPES, pH 7.5). The pellet was subsequently suspended in lysis buffer (5 mL, 100 mM HEPES, pH 7.5 containing 0.2 mM PLP) and lysed by ultrasonication (on ice, 12 cycles of 10 seconds on/ 10 seconds off, duty cycle 6, Branson Sonifier). Cellular debris was pelleted by centrifugation (15 min, 12000 rpm, 4°C) and the supernatant was stored in aliquots (550 µL, -20 °C). Overexpression was confirmed by SDS-PAGE (4-20% precast gradient gel, Bis-Tris running buffer, 200V, 40 min). Protein concentrations were measured with Bradford reagent in 1 mL cuvettes at 595 nm and BSA as external standard.

Colorimetric 4-nitrophenyl-2-ethylamine transaminase assay: h*olo*-transaminases (0.4-3.3 mg/mL total protein, 0.5 mM PLP) were mixed with ketoacid substrate (10 mM, 1.0 eq.) and the reaction was initiated by the addition of 4-nitrophenyl-2-ethylamine·HCl in 96-well plates (25 mM, 2.5 eq., 200 µL final volume, 30°C, 500 rpm, 15 min – 24 h, 100 mM HEPES, pH 7.5). 0-40% (v/v) DMSO was added as co-solvent in some cases. Negative control: empty vector lysates.

Preparation of D-ketogluconate stock solutions: solutions of D-ketogluconate hemicalcium were mixed with sodium phosphate buffer (1.0 eq.). Precipitated calcium phosphate was removed by centrifugation (12000 rpm, 5 min.).

Solubility test of PLP under reaction conditions: solutions of PLP (0.2, 0.5 or 1 mM, in 100 mM HEPES, pH 7.5) were mixed with calcium chloride (5, 10, 15, 25, 50, 75, 90 mM, 200 µL final volume) and incubated in 96-well plates at room temperature for 24 hours. Possible precipitation was inspected visually.

Methylbenzylamine (MBA) assay: reactions with (*R*) or (*S*) MBA as amine donor were carried out in 1.5 mL Eppendorf tubes in a final volume of 1.5 mL. *Holo*-transaminases (1-3 mg/mL lysate, 0.5 mM PLP) were mixed with ketoacid substrate (10 mM, 1.0 eq.) and the reaction was initiated by the addition of the appropriate enantiomer of MBA (25 mM, 2.5 eq., 30°C, 500 rpm). Conversions were determined indirectly *via* the concentration of acetophenone (sampling: 150 µL + 450 µL 0.4% (v/v) TFA, 13200 rpm, 10 minutes, 4°C, RP-HPLC: ACE 5 C18-AR column, 150x4.6 mm on a 1260 Infinity, Agilent Technologies, 1 mL/min, 25°C, 254 nm, gradient method: isocratic (1 minute): 15% MeCN, 85% 0.1% (v/v) TFA followed by a gradient going to 72% MeCN within 9 minutes and subsequent isocratic equilibration for 5 minutes. Retention time acetophenone (8.9 minutes), with quantitation *via* an external standard methodology, R^2 = 0.995). For inhibition studies with transaminase pQR2189, calcium and phosphate were added in the corresponding amounts before the addition of the substrates, but after holoenzyme formation.

Influence of DMSO on equilibrium concentrations of D-ketogluconate: D-ketogluconate hemicalcium was dissolved in D_2O (150 mM) and was supplemented with DMSO-D_6 (0%, 10%, 20% v/v). The relative integrals of the pyranose, furanose and open chain forms were compared to analyse the influence of DMSO on the equilibrium constants.

Table 1: Overview of transaminases screened in the first round.

pQR	UniProt ID	Microorganism	Class	Selectivity	kDa	Reference
801	Q7NWG4	*Chromobacterium violaceum*	III	*S*	51.2	[30]
1005	A6T537	*Klebsiella pneumoniae*	III	*S*	45.5	[31, 43]
1019	Q3IWE9	*Rhodobacter sphaeroides*	III	*S*	50.1	[31, 43]
1756	Q59228	*Geobacillus stearothermophilus*	I	*S*	42.4	[31]
1757	Q1IZC2	*Deinococcus geothermalis*	V	*S*	40.9	[31]
2189	MK121625	*Metagenomic*	III	*S*	52.3	[2]
2191	MK121627	*Metagenomic*	III	*S*	49.2	[2]
2208	MK121644	*Metagenomic*	III	*S*	50.1	[2]
4516	A1TDP1	*Mycobacterium vanbaalenii*	IV	*R*	36.6	[3, 44]

Acetone tolerance of TAms using the colorimetric assay: acetone tolerance of transaminases was investigated using an adapted version of the the colorimetric assay. Lysate (10 µL, 0.5 – 1 mg/mL total protein), PLP (0.25 mM) and sodium pyruvate (10 mM, 1.0 eq.) were mixed with acetone (0 - 40% v/v) and the reaction was initiated by the addition of 4-nitrophenylethylamine (25 mM, 2.5 eq., 6 h, 30°C, 200 rpm). Negative controls: no enzyme, 40% acetone. Positive controls: sodium pyruvate, 0% acetone.

Transamination of D-glucosaminic acid with acetone: D-Glucosaminic acid (10 mM, 1.0 eq.) was added to a solution of lysate (5 mg/mL) and PLP (0.5 mM) in KPi buffer (100 mM, pH 7.5) and the reaction was initiated by the addition of acetone (250 mM, 25.0 eq., 30°C, 300 rpm, 18-24 hours) and subsequently quenched/derivatised with Fmoc chloride according to the protocol provided below. Negative controls: empty plasmid cell free extracts and reaction mixtures without enzyme.

Fmoc-derivatisation: the sample (250 µL), was diluted with borate buffer (250 µL, 200 mM, pH 9.0) and fluorenylmethyloxycarbonyl chloride (Fmoc Cl, 40 mM, 4.0 eq., in acetonitrile, 250 µL) was added. The mixture was incubated at room temperature (10 minutes) and excess Fmoc Cl was quenched by the addition of L-alanine (250 µL, 100 mM, 10.0 eq., in 200 mM borate buffer, pH 9.0, 10 minutes). MeOH was added (500 µL) and precipitate was removed by centrifugation. Samples were analysed by RP-HPLC (Ace-5 C18) using a gradient method (5-95% MeCN/H_2O over 30 minutes) with detection at 214 nm. Retention times: D-glucosaminic acid-Fmoc: 15.8 min., IPA-Fmoc: 24. 7 min.

Derivatisation was confirmed for **IPA-Fmoc** by LC-MS and NMR. M/z predicted: 282.2, found: 282.4 m/z. ^{1}H-NMR (700 MHz, $CDCl_3$) δ 8.29 (s, 2H, H3), 7.76 (d, *J* = 7.5 Hz, 2H, H11), 7.59 (dd, *J* = 7.5, 0.8 Hz, 2H, H8), 7.42 – 7.38 (m, 2H, H10), 7.33 – 7.30 (m, 2H, H9), 4.44 – 4.37 (m, 2H, H5), 4.21 (t, *J* = 6.7 Hz, 1H, H6), 3.51 (dt, *J* = 12.3, 6.1 Hz, 1H, H2), 1.43 (d, *J* = 6.6 Hz, 6H, H1). ^{13}C-NMR (176 MHz, $CDCl_3$, ^{1}H decoupled) δ 155.71 (C4), 144.18 (C7), 141.45 (C12), 127.77 (C10), 127.14 (C9), 125.16 (C8), 120.08 (C11), 66.47 (C5), 47.47 (C6), 44.73 (C2), 21.06 (C1). LC-MS (D-glucosaminic acid-Fmoc), m/z predicted: 418.2, found: 418.5.

NMR of D-glucosaminic acid: ^{1}H-NMR (700 MHz, Deuterium Oxide) δ 4.45 (dd, *J* = 3.7, 1.4 Hz, 1H, H3), 3.91 (d, *J* = 3.7 Hz, 1H, H2), 3.85 (dd, *J* = 11.9, 2.6 Hz, 1H, H_a6), 3.78 (m, 1H, H4), 3.72 (m, 1H, H5), 3.68 (dd, *J* = 11.9, 5.5 Hz, 1H, H_b6).^{13}C-NMR (176 MHz, D_2O, ^{1}H decoupled) δ 173.32 (C1), 73.34 (C4), 71.31 (C5), 67.81 (C3), 63.32 (C6), 58.91 (C2).

Expression of pQR2501-2594 in 96-well plates: 94 thermostable transaminases with pQR numbers 2501-2594, pQR2025 (*Chromobacterium violaceum* transaminase) and an empty plasmid as negative control were grown in a 96 deep-well plate overnight (37°C, 1200 rpm, 1 mL TB medium, 50 µg/mL kanamycin, gas permeable membrane seals, *E.coli* BL21(DE3)). From this, 10 µL were used as inoculum for the expression medium (1 mL, 0.5 M sorbitol, 0.1 mM PLP, 100 mM NaPi, 1x TB medium, 50 µg/mL kanamycin, 37°C, 1200 rpm). The cells were grown until an OD_{600} of 0.6-1 was reached and gene expression was induced by the addition of IPTG (0.5 mM, 20°C, 1200 rpm, 16 h). The cells were pelleted by centrifugation (4°C, 20 min., 3900 rpm) and stored at -20°C. The cells were subsequently suspended in buffer (100 mM NaPi, pH 7.0, 0.1 mM PLP, 200 µL, 1200 rpm, 10 minutes) and lysed by ultrasonication (on ice, 20 minutes), rotating the plate by 90° every three minutes. The lysate was transferred into a 96-well plate and the debris was pelleted by centrifugation (4°C, 20 min., 3900 rpm).

DL-serine/hydroxypyruvate activity assay: conversion of DL-serine/hydroxypyruvate as an amine donor/acceptor pair was investigated for pQR2501-2594 by the previously developed colorimetric assay with WST-1.[38] DL-serine (10 mM, 1.0 eq.), PLP (0.1 mM) and sodium pyruvate (10 mM, 1.0 eq.) were incubated with 20 µL of lysate (100 mM NaPi, pH 7.0, 200 µL final volume, 50°C, 18 h). The reaction mixture (20 µL) was subsequently mixed with WST-1 (180 µL, 0.1 mg/mL in 0.1 M NaOH, 10 minutes, room temperature). The absorbance was measured at 600 nm and quantified using a calibration line with hydroxypyruvate as external standard. Negative control: empty vector plasmid.

Separation of D-glucosaminic acid and DL-serine by ion chromatography: amino acids were separated by anion exchange chromatography using a Dionex Aminopac PA1 anion exchange column (4 x 250 mm, fitted with a Dionex Aminopac PA1 guard column 4 x 50 mm, 100 mM KOH, 0.25 mL/min, 30°C) using electrochemical detection. Retention times: D-glucosaminic acid: 9.8 min., DL-serine: 4.9 min.

96-well activity screening for the conversion of D-ketogluconate: DL-serine (20 mM, 1.0 eq. per enantiomer.), PLP (0.1 mM) and D-ketogluconate (10 mM, 1.0 eq.) were incubated with 20 µL of lysate (100 mM NaPi, pH 7.0, 200 µL final volume, 50°C, 18 h). The D-ketogluconate hemicalcium stock solution was prepared in milliQ and the cation was exchanged by precipitation with NaPi buffer. The reaction was quenched by the addition of TFA (0.5% v/v, 1:1 dilution) and precipitates were removed by centrifugation (4°C, 20 min., 3900 rpm). The crude mixture was diluted with milliQ (100x) and analysed by ion chromatography.

96-well activity screening for the conversion of D-glucosaminic acid: hydroxypyruvate (10 mM, 1.0 eq.), PLP (0.1 mM) and D-glucosaminic acid (10 mM, 1.0 eq.) were incubated with 20 µL of lysate (100 mM NaPi, pH 7.0, 200 µL final volume, 50°C, 18 h). The reaction was quenched by the addition of TFA (0.5% v/v, 1:1 dilution) and precipitates were removed by centrifugation (4°C, 20 min., 3900 rpm). The crude mixture was diluted with milliQ (100x) and analysed by ion chromatography.

SDS-PAGE analysis

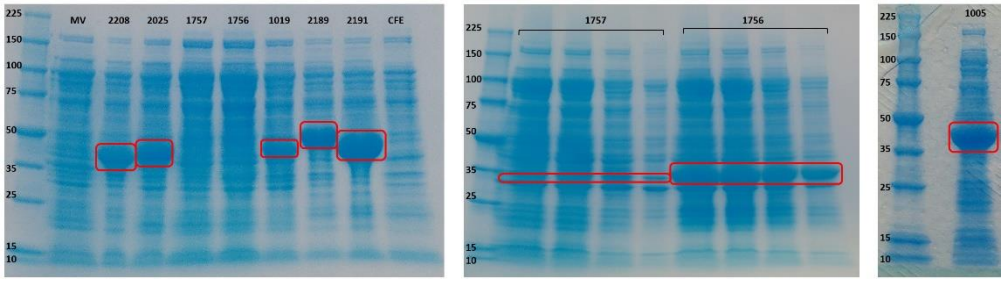


Figure 7: SDS-PAGE analysis of overexpressed transaminases. Molecular weights (kDa): MV (36.6), 2208 (50.1), 2025 (51.2), 1757 (40.9), 1756 (42.4), 1019 (50.1), 2189 (52.3), 2191 (49.2), 1005 (45.5). Extended denaturation in boiling SDS sample buffer (>10 minutes) was required to denature the homodimer of thermostable transaminases pQR1757 (*Deinococcus Geothermalis*) and pQR1756 (*Geobacillus Stearothermophilus*).

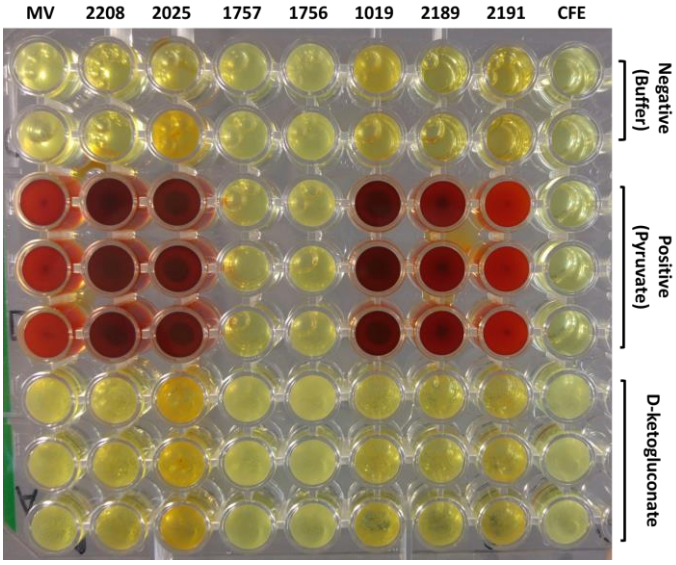


Figure 8: Colorimetric screening for conversion of D-ketogluconate (0.4-0.65 mg/mL total protein). pQR1756 and 1757 do not accept 4-phenylethylamine as amine donor. The negative control ("CFE") did not show any endogenous background activity.

NMR of D-ketogluconate hemicalcium: Major pyranose: ^{1}H-NMR (700 MHz, Deuterium Oxide) δ 3.85 (dd, *J* = 10.2, 3.4 Hz, H4), 3.76 (dd, *J* = 13.0, 2.1 Hz, H6a), 4.06-4.0 (m, H3, H5, H6b). ^{13}C-NMR (176 MHz, D_2O, ^{1}H decoupled) δ 175.43 (C1), 97.83 (C2), 70.61 (C4), 69.96 (C3), 69.95 (C5), 64.98 (C6).

Minor furanose: ^{1}H-NMR (700 MHz, Deuterium Oxide) δ 4.31 (d, *J* = 8.2 Hz, H3'), 4.07 (d, *J* = 8.0 Hz, H4'), 3.99 – 3.95 (m, H5'), 3.81 (dd, *J* = 12.4, 3.1 Hz, H6a'), 3.70 (dd, *J* = 12.4, 5.7 Hz, H6b'). ^{13}C-NMR (176 MHz, D_2O, ^{1}H decoupled) δ 176.09 (C1'), 100.33 (C2'), 81.66, (C5'), 78.93 (C3'), 74.89 (C4'), 62.63 (C6').

Linear ketone/hydrate forms: ^{1}H-NMR (700 MHz, Deuterium Oxide) δ 4.18-4.12 (H4'', H4''', H5'', H5'''), 3.91-3.89 (H6'', H6'''), 3.83-3.80 (H3'', H3'''). ^{13}C-NMR (176 MHz, D_2O, ^{1}H decoupled) δ 175.23, 174.78 (C1'' & C1'''), 105.09 (C2'''), 97.66, 72.83 (C3'' & C3'''), 83.49, 82.67, 76.32, 71.72 (C4'', C4''', C5'', C5'''), 65.08, 61.58 (C6'', C6''').

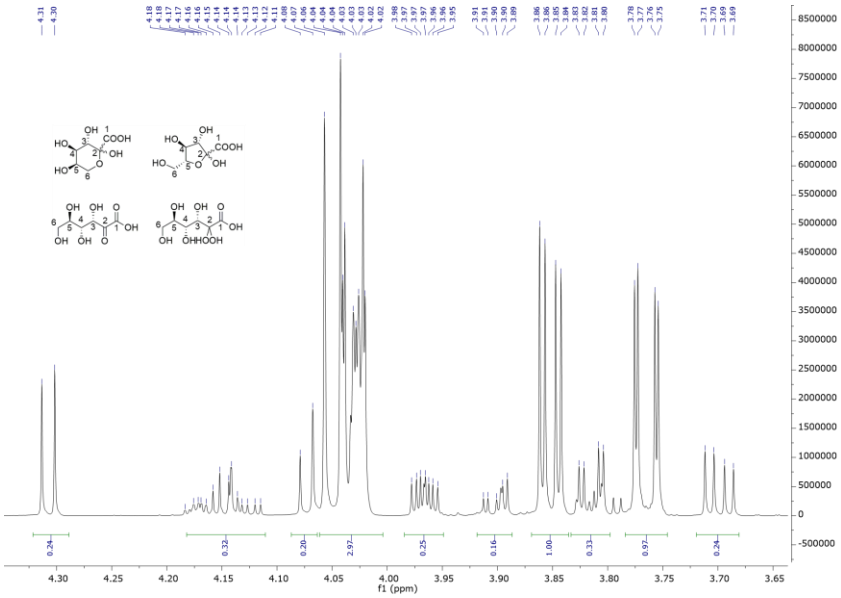


Figure 9: ^{1}H-NMR of D-ketogluconate hemicalcium salt in D_2O.

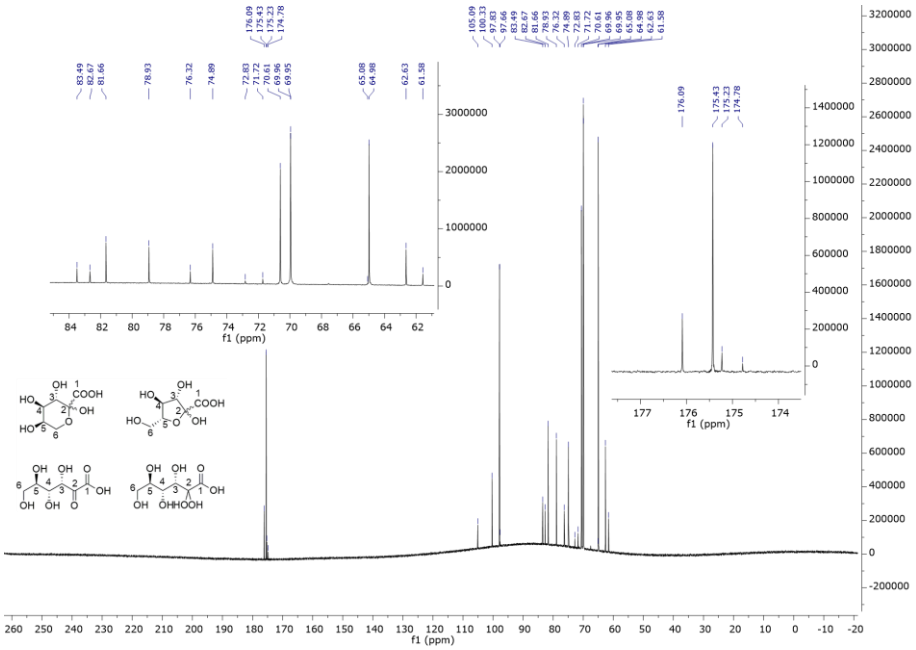


Figure 10: ^{13}C-NMR of D-ketogluconate hemicalcium salt in D_2O.

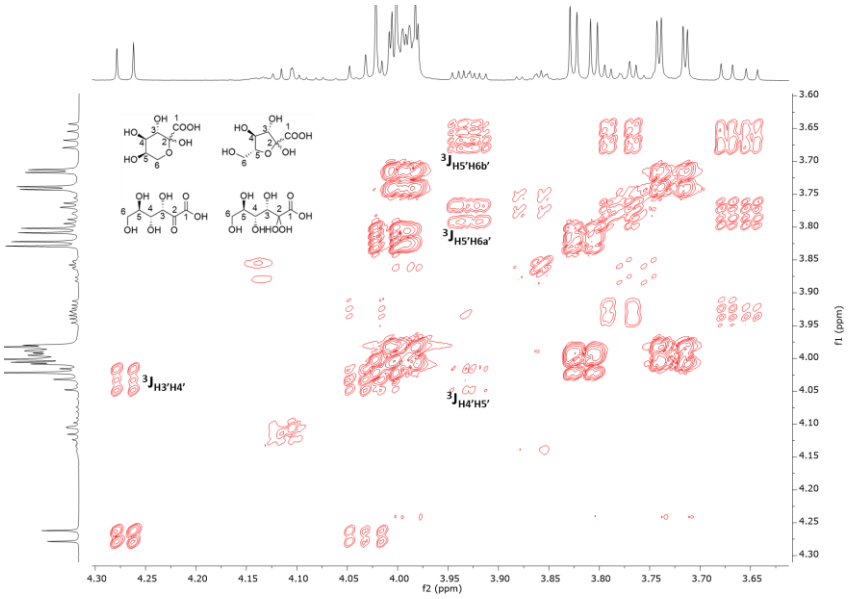


Figure 11: $^{1}H^{1}H$ COSY spectrum of D-ketogluconate hemicalcium salt in D_2O.

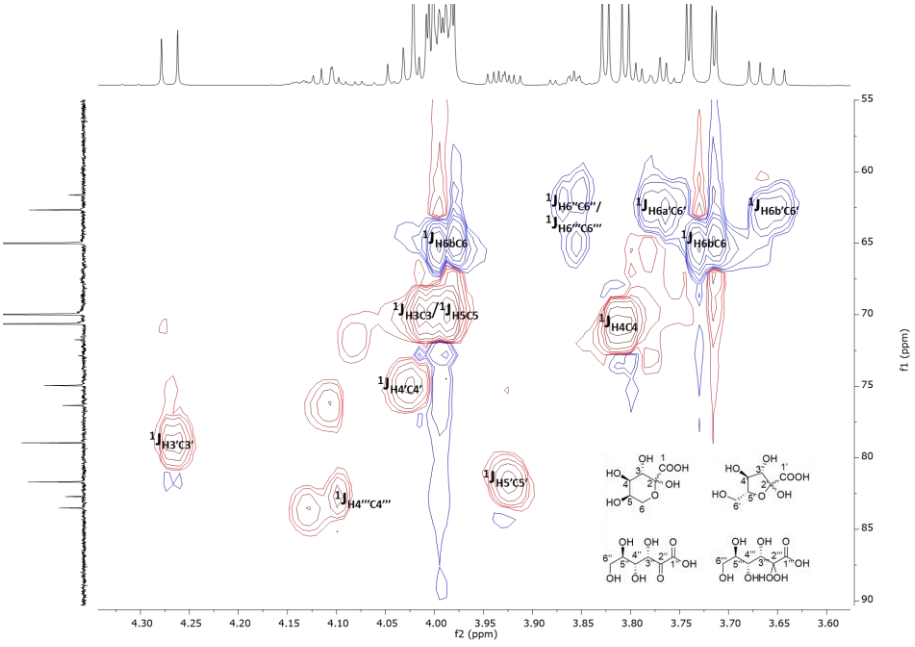


Figure 12: HSQC spectrum of D-ketogluconate hemicalcium salt in D_2O.

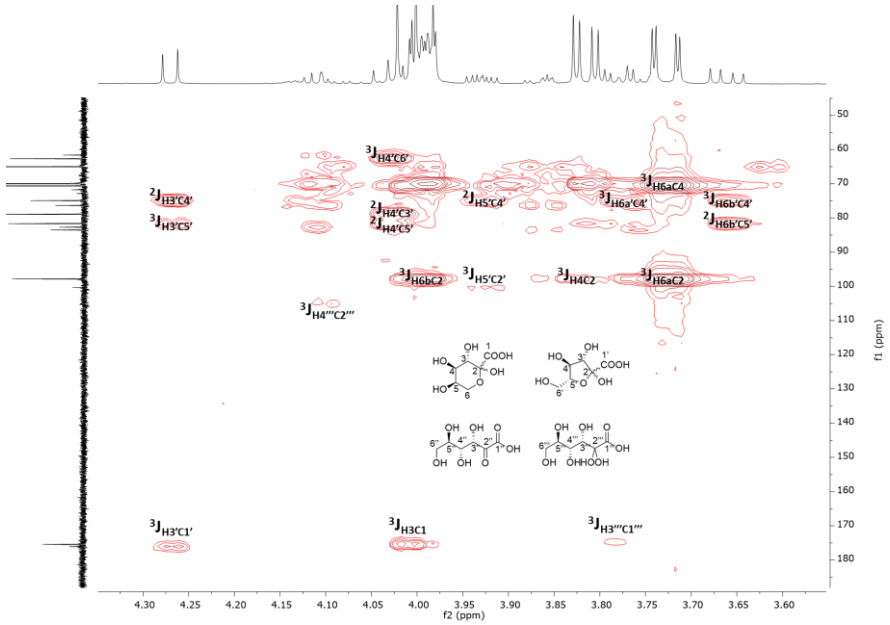


Figure 13: HMBC spectrum of D-ketogluconate hemicalcium in D_2O.

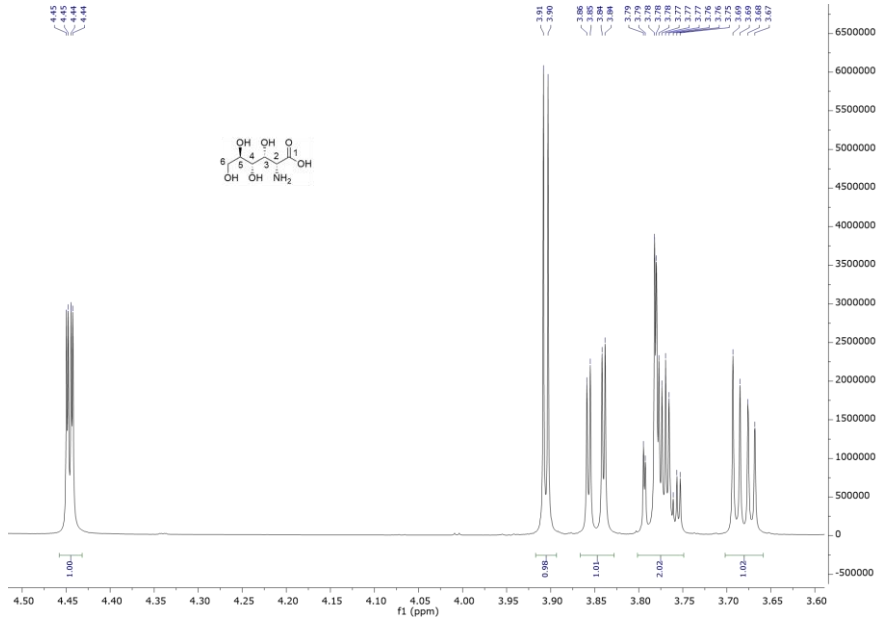


Figure 14: ^{1}H-NMR of D-glucosaminic acid in D_2O.

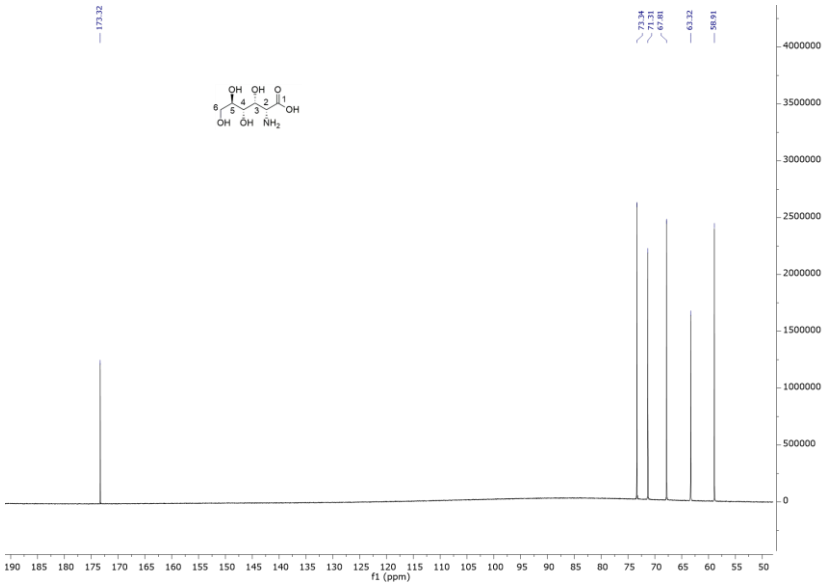


Figure 15: ^{13}C-NMR of D-glucosaminic acid in D_2O.

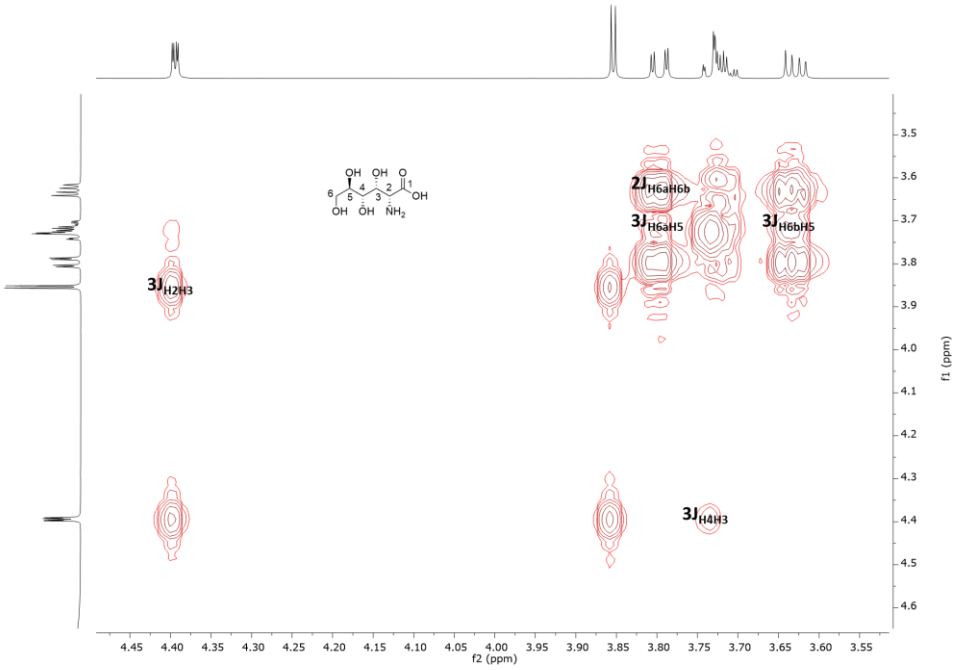


Figure 16: ^{1}H^{1}H COSY spectrum of D-glucosaminic acid in D_2O.

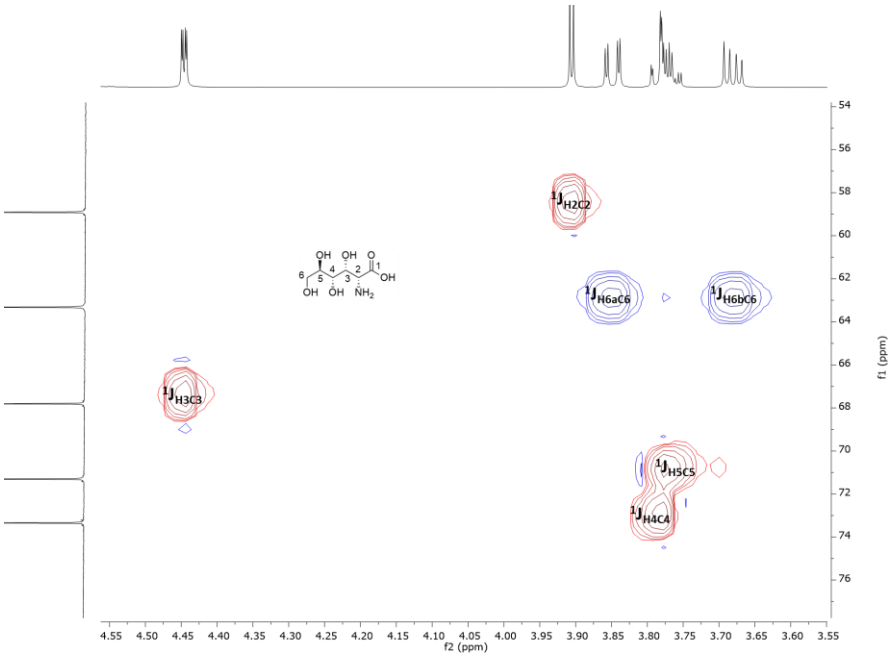


Figure 17: HSQC spectrum of D-glucosaminic acid in D_2O.

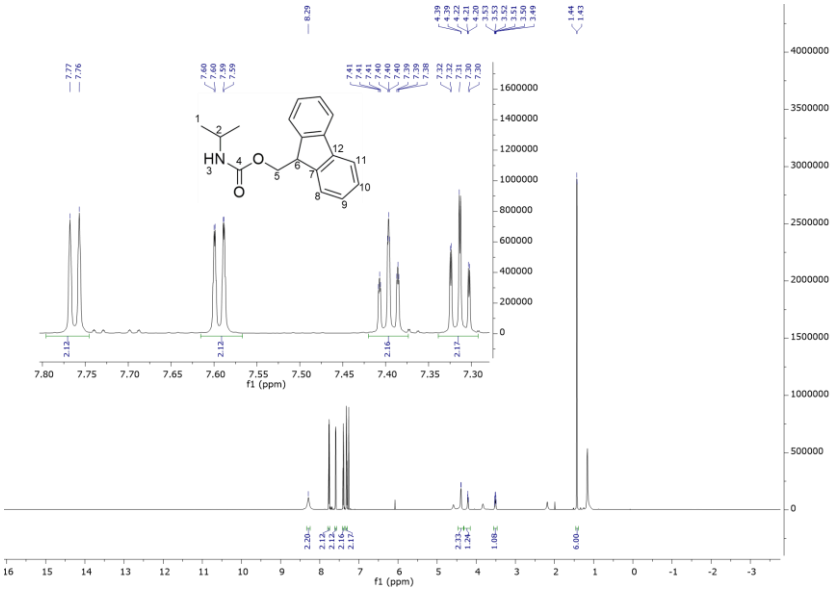


Figure 18: ^{1}H-NMR spectrum of IPA-Fmoc in $CDCl_3$.

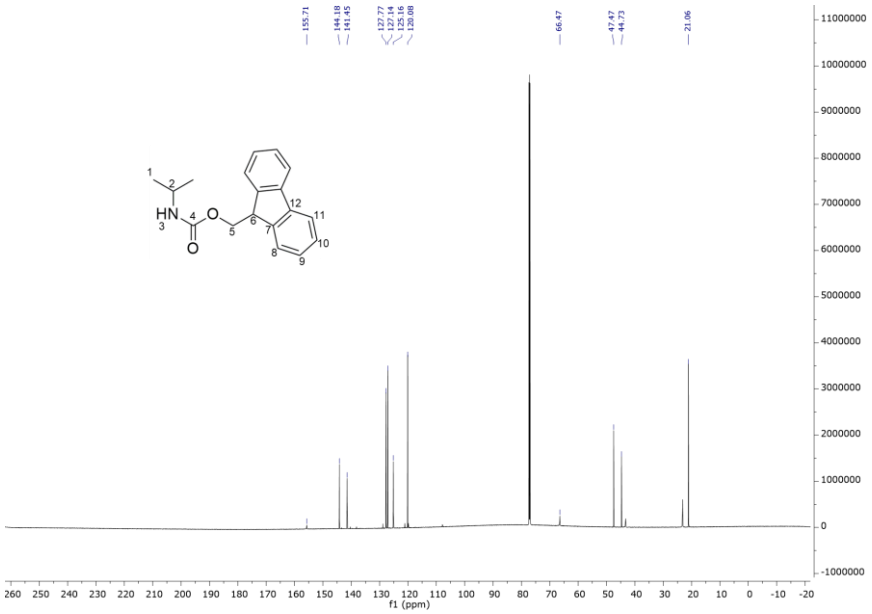


Figure 19: ^{13}C-NMR spectrum of IPA-Fmoc in $CDCl_3$.

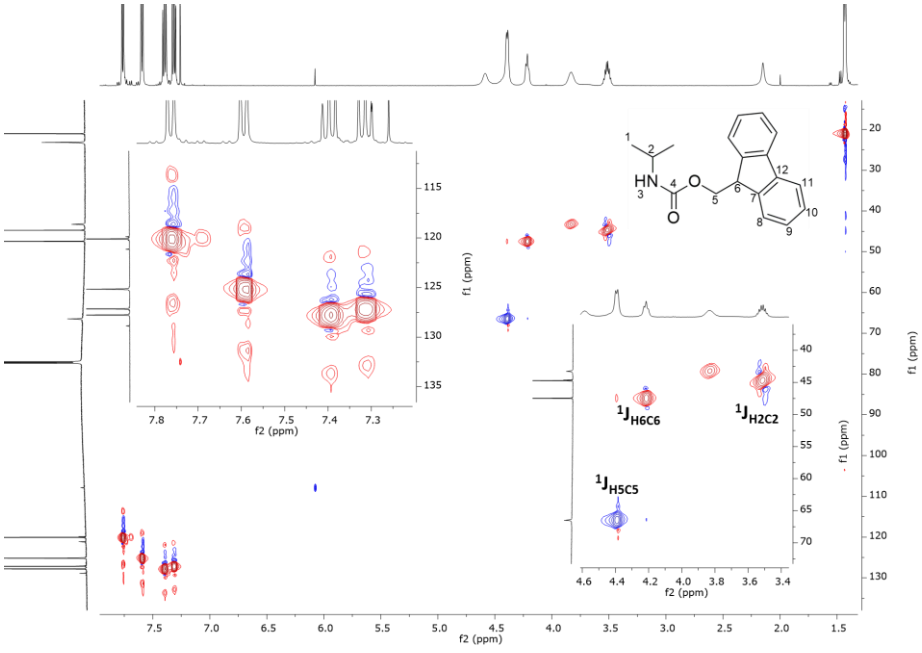


Figure 20: HMBC spectrum of IPA-Fmoc in $CDCl_3$.

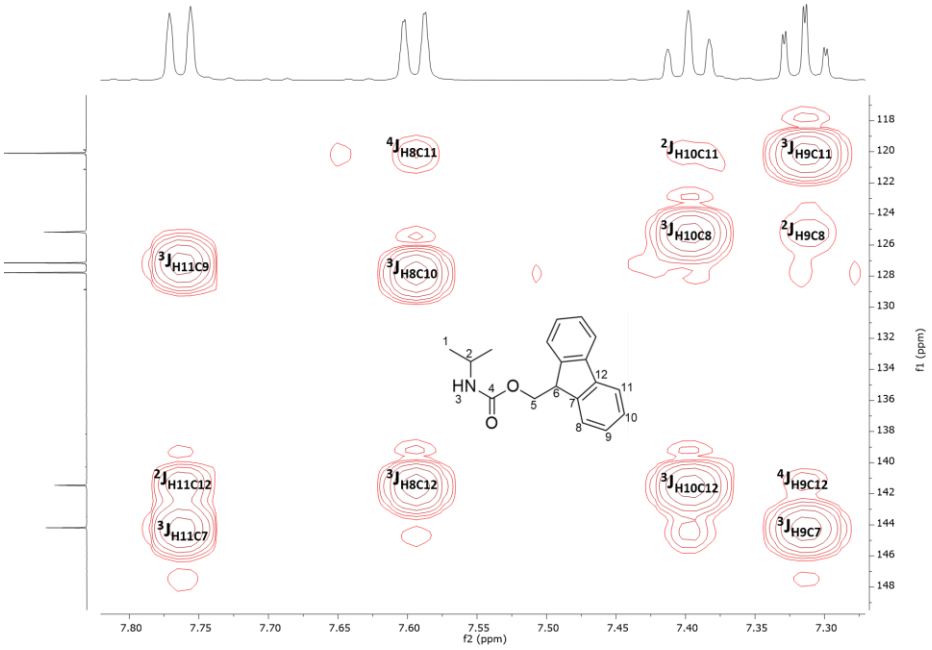


Figure 21: HSQC spectrums of IPA-Fmoc in $CDCl_3$.

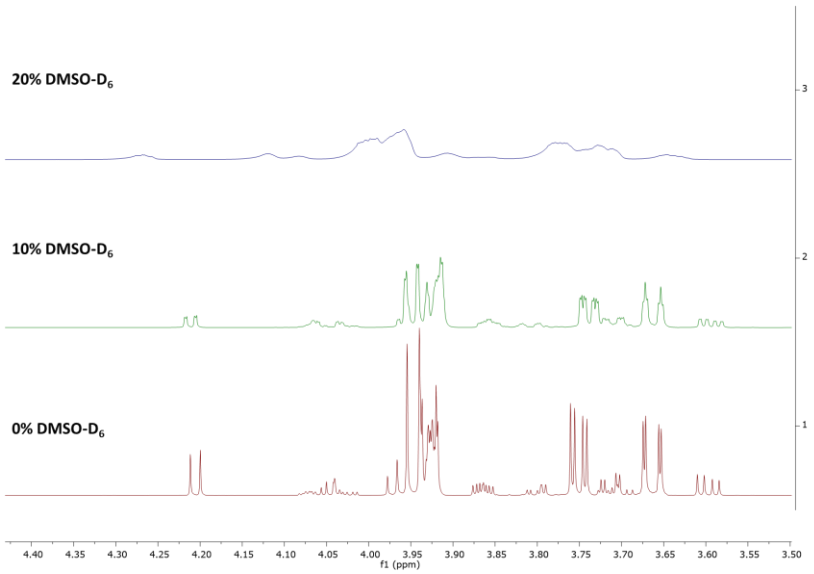


Figure 22: ^{1}H-NMR spectrum of D-ketogluconate in D_2O with 0-20% DMSO-D_6.

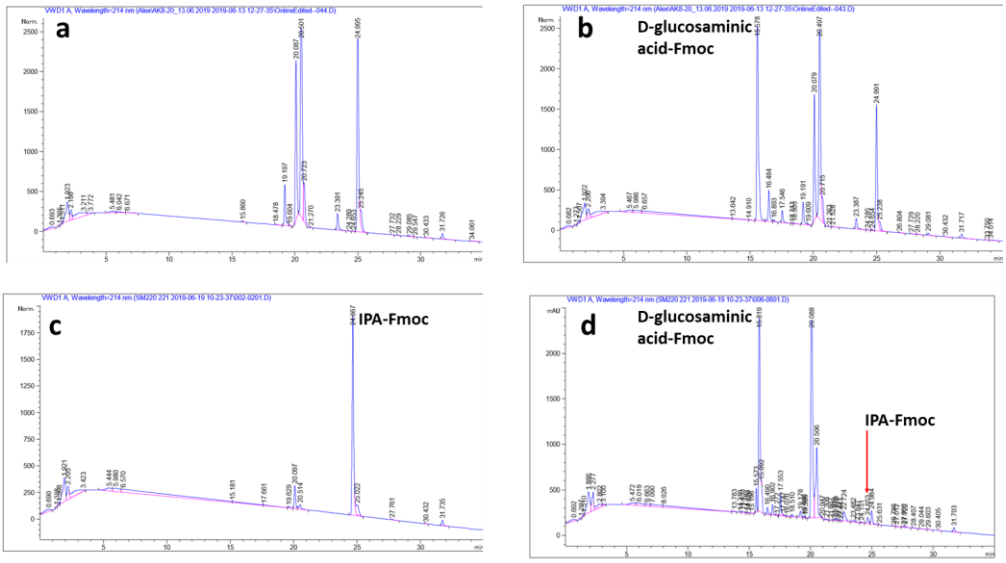


Figure 23: (a) Quenching of Fmoc-Cl with L-alanine in MeCN/MeOH to yield L-ala-Fmoc, Fmoc-OH and Fmoc-OMe. **(b)** Derivatisation of D-glucosaminic acid with Fmoc-Cl (15.58 min), quenched by the addition of L-alanine. **(c)** Derivatisation of IPA with Fmoc-Cl (24.67 min). **(d)** Transamination of D-glucosaminic acid with acetone. Formation of isopropylamine (detection as IPA-Fmoc) is indicated by a peak at 24.65 minutes and suggests conversion. No peak was observed at 24.6 minutes in the absence of either enzyme or D-glucosaminic acid.

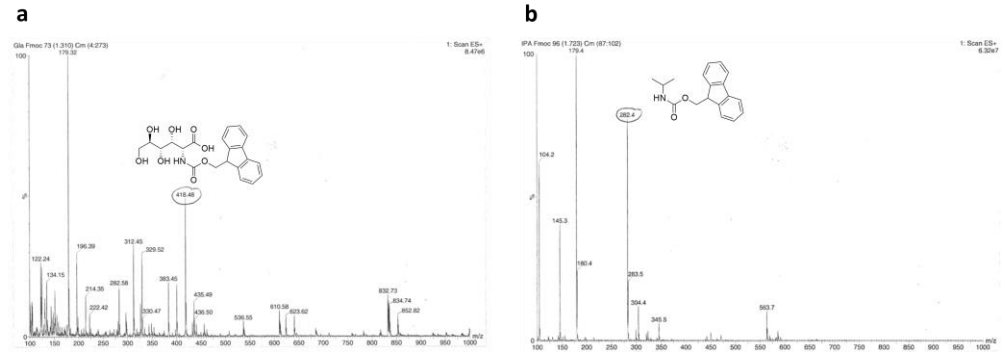


Figure 24: (a) LC-MS spectrum of D-glucosaminic acid-Fmoc. Predicted: 418.15 m/z, found: 418.48 m/z. **(b)** LC-MS spectrum of IPA-Fmoc. Predicted: 282.15 m/z, found: 282.4 m/z.

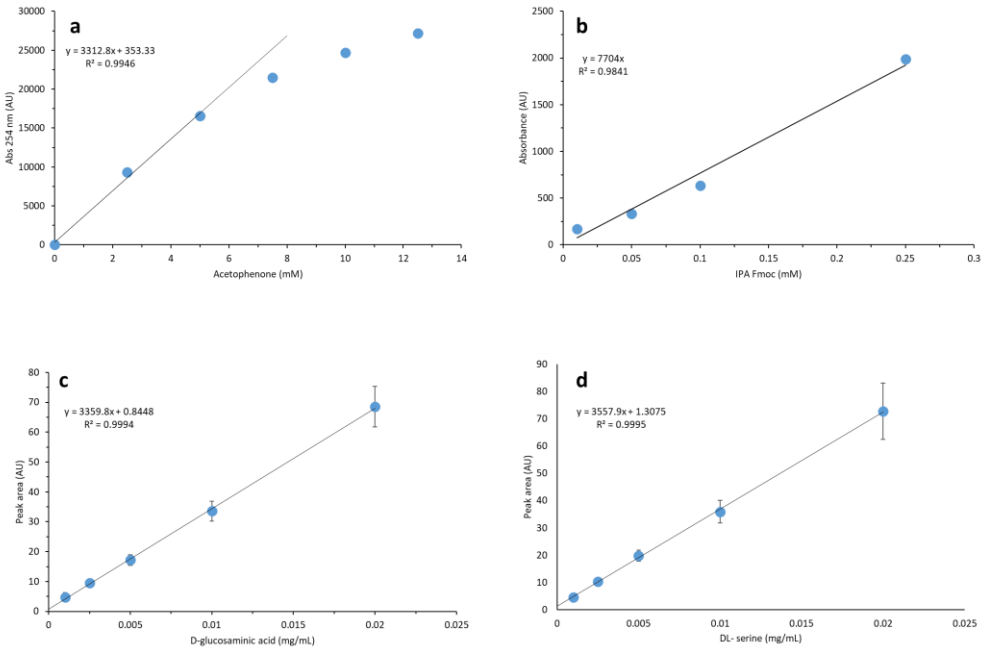


Figure 25: (a) Calibration curve for acetophenone (RP-HPLC, C18, 254 nm). The linear range only extended up to 5 mM. Dilutions were prepared accordingly to remain within the linear range during analysis. **(b)** Calibration curve of IPA-Fmoc (RP-HPLC, C18, 214 nm). **(c)** D-glucosaminic acid calibration curve on the ICS system. **(d)** DL-serine calibration curve on the ICS system.

References

[1] Rosenthaler, L., Durch enzyme bewirkte asymmetrische synthesen. *Biochem. Z*. **1908**, *14*, 238-253.

[2] Leipold, L.; Dobrijevic, D.; Jeffries, J. W.; Bawn, M.; Moody, T. S.; Ward, J. M.; Hailes, H. C., The identification and use of robust transaminases from a domestic drain metagenome. *Green Chemistry,* **2019**, *21* (1), 75-86.

[3] Höhne, M.; Schätzle, S.; Jochens, H.; Robins, K.; Bornscheuer, U. T., Rational assignment of key motifs for function guides in silico enzyme identification. *Nature Chemical Biology,* **2010**, *6* (11), 807.

[4] van den Bergh, T.; Tamo, G.; Nobili, A.; Tao, Y.; Tan, T.; Bornscheuer, U. T.; Kuipers, R. K.; Vroling, B.; de Jong, R. M.; Subramanian, K., CorNet: Assigning function to networks of co-evolving residues by automated literature mining. *PloS One*, **2017**, *12* (5), e0176427.

[5] Arnold, F. H., Design by directed evolution. *Accounts of Chemical Research*, **1998**, *31*, (3), 125-131.

[6] Arnold, F. H., Directed evolution: bringing new chemistry to life. *Angewandte Chemie International Edition,* **2018**, *57* (16), 4143-4148.

[7] Renata, H.; Wang, Z. J.; Arnold, F. H., Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angewandte Chemie International Edition,* **2015**, *54,* (11), 3351-3367.

[8] Romero, P. A.; Arnold, F. H., Exploring protein fitness landscapes by directed evolution. *Nature Reviews Molecular Cell Biology,* **2009**, *10* (12), 866.

[9] Jäckel, C.; Kast, P.; Hilvert, D., Protein design by directed evolution. *Annual Reviews of Biophysics*, **2008**, *37*, 153-173.

[10] Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; Clair, J. L. S.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L., Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science,* **2010**, *329* (5989), 309-313.

[11] Turner, N. J.; O'reilly, E., Biocatalytic retrosynthesis. *Nature Chemical Biology*, **2013**, *9* (5), 285.

[12] Drienovská, I.; Roelfes, G., Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nature Catalysis,* **2020**, *3*, 193-202.

[13] Hohsaka, T.; Sisido, M., Incorporation of non-natural amino acids into proteins. *Current Opinion in Chemical Biology*, **2002**, *6* (6), 809-815.

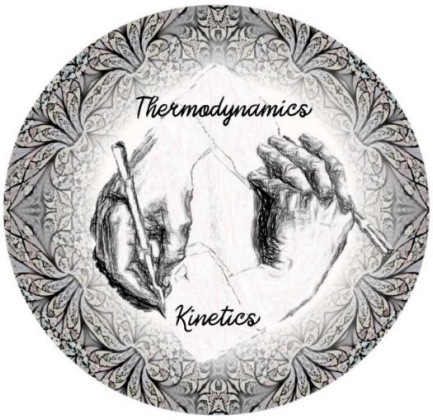
[14] Hohsaka, T.; Ashizuka, Y.; Taira, H.; Murakami, H.; Sisido, M., Incorporation of nonnatural amino acids into proteins by using various four-base codons in an *Escherichia coli in vitro* translation system. *Biochemistry,* **2001**, *40* (37), 11060-11064.

[15] Sisido, M.; Hohsaka, T., Introduction of specialty functions by the position-specific incorporation of nonnatural amino acids into proteins through four-base codon/anticodon pairs. *Applied Microbiology and Biotechnology,* **2001**, *57* (3), 274-281.

[16] Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M., Five-base codons for incorporation of nonnatural amino acids into proteins. *Nucleic Acids Research,* **2001**, *29* (17), 3646-3651.
[17] Marsden, S. R.; Gjonaj, L.; Eustace, S. J.; Hanefeld, U., Separating Thermodynamics from Kinetics—A New Understanding of the Transketolase Reaction. *ChemCatChem,* **2017**, *9* (10), 1808-1814.
[18] Marsden, S.; Mestrom, L.; Bento, I.; McMillan, D.; Hagedoorn, P.-L.; Hanefeld, U., CH-π Interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase. *Advanced Synthesis & Catalysis,* **2019**, *361*, 2649-2658.
[19] Carter, P.; Wells, J. A., Dissecting the catalytic triad of a serine protease. *Nature,* **1988**, *332* (6164), 564.
[20] Dodson, G.; Wlodawer, A., Catalytic triads and their relatives. *Trends in Biochemical Sciences,* **1998**, *23* (9), 347-352.
[21] Skarbek, K.; Milewska, M. J., Biosynthetic and synthetic access to amino sugars. *Carbohydrate Research,* **2016**, *434*, 44-71.
[22] Wen-xiu, G.; Wen-shui, X., Catalytic Synthesis of D-Glucosaminic Acid from D-Glucosamine on Active Charcoal-Supported Pd-Bi Catalysts. *Journal of Carbohydrate Chemistry,* **2006**, *25* (4), 297-301.
[23] Pezzotti, F.; Therisod, H.; Therisod, M., Enzymatic synthesis of D-glucosaminic acid from D-glucosamine. *Carbohydrate Research,* **2005**, *340* (1), 139-141.
[24] Wu, B.; Bai, Z.; Meng, X.; He, B., Efficient production of D-glucosaminic acid from D-glucosamine by Pseudomonas putida GNA5. Biotechnology progress 2011, 27 (1), 32-37.

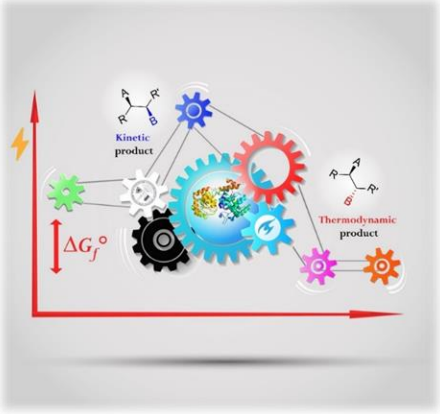
[25] Kim, J.; Kershner, J. P.; Novikov, Y.; Shoemaker, R. K.; Copley, S. D., Three serendipitous pathways in E. coli can bypass a block in pyridoxal-5′-phosphate synthesis. *Molecular Systems Biology,* **2010**, *6* (1), 436.
[26] Steinreiber, J.; Fesko, K.; Reisinger, C.; Schürmann, M.; van Assema, F.; Wolberg, M.; Mink, D.; Griengl, H., Threonine aldolases—an emerging tool for organic synthesis. *Tetrahedron,* **2007**, *63* (4), 918-926.
[27] Steinreiber, J.; Fesko, K.; Mayer, C.; Reisinger, C.; Schürmann, M.; Griengl, H., Synthesis of γ-halogenated and long-chain β-hydroxy-α-amino acids and 2-amino-1, 3-diols using threonine aldolases. *Tetrahedron,* **2007**, *63* (34), 8088-8093.
[28] Fesko, K.; Uhl, M.; Steinreiber, J.; Gruber, K.; Griengl, H., Biocatalytic access to a, a-dialkyl-a-amino acids by a mechanims-based approach. *Angewandte Chemie International Edition,* **2010**, *49*, 121-124.
[29] Fesko, K.; Reisinger, C.; Steinreiber, J.; Weber, H.; Schürmann, M.; Griengl, H., Four types of threonine aldolases: similarities and differences in kinetics/thermodynamics. *Journal of Molecular Catalysis B: Enzymatic,* **2008**, *52*, 19-26.
[30] Deszcz, D.; Affaticati, P.; Ladkau, N.; Gegel, A.; Ward, J. M.; Hailes, H. C.; Dalby, P. A., Single active-site mutants are sufficient to enhance serine: pyruvate α-transaminase activity in an ω-transaminase. *The FEBS Journal,* **2015**, *282* (13), 2512-2526.
[31] Bawn, M.; Subrizi, F.; Lye, G. J.; Sheppard, T. D.; Hailes, H. C.; Ward, J. M., One-pot, two-step transaminase and transketolase synthesis of L-gluco-heptulose from L-

arabinose. *Enzyme and Microbial Technology,* **2018**, *116*, 16-22.
[32] Hernandez, K.; Bujons, J.; Joglar, J.; Charnock, S. J.; Dominguez de Maria, P.; Fessner, W. D.; Clapés, P., Combining Aldolases and Transaminases for the Synthesis of 2-Amino-4-hydroxybutanoic Acid. *ACS Catalysis,* **2017**, *7* (3), 1707-1711.
[33] De Berardinis, V.; Guérard-Hélaine, C.; Darii, E.; Bastard, K.; Hélaine, V.; Mariage, A.; Petit, J.-L.; Poupard, N.; Sánchez-Moreno, I.; Stam, M., Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chemistry,* **2017**, *19* (2), 519-526.
[34] Laurent, V.; Uzel, A.; Hélaine, V.; Nauton, L.; Traikia, M.; Gefflaut, T.; Salanoubat, M.; de Berardinis, V.; Lemaire, M.; Helaine, C., Exploration of Aldol Reactions Catalysed by Stereoselective Pyruvate Aldolases with 2-Oxobutyric Acid as Nucleophile. *Advanced Synthesis & Catalysis,* **2019**, *361*, 2713-2717.
[35] Fang, J.; Hait, D.; Head-Gordon, M.; Chang, M., Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angewandte Chemie International Edition,* **2019**, *131*, 11967-11971.
[36] Laurent, V.; Gourbeyre, L.; Uzel, A.; Hélaine, V.; Nauton, L.; Traïkia, M.; De Berardinis, V.; Salanoubat, M.; Gefflaut, T.; Lemaire, M., Pyruvate Aldolases Catalyse Cross-Aldol Reactions between Ketones: Highly Selective Access to Multi-Functionalized Tertiary Alcohols. *ACS Catalysis,* **2020**, *10* (4), 2538-2543.
[37] Good, N. E.; Winget, G. D.; Winter, W.; Connolly, T. N.; Izawa, S.; Singh, R. M., Hydrogen ion buffers for biological research. *Biochemistry,* **1966**, *5*, 467-477.
[38] Bommer, M.; Ward, J. M., Micromolar colorimetric detection of 2-hydroxy ketones with the water-soluble tetrazolium WST-1. *Analytical Biochemistry,* **2016**, *493*, 8-10.
[39] Drewke, C.; Klein, M.; Clade, D.; Arenz, A.; Müller, R.; Leistner, E., 4-O-phosphoryl-L-threonine, a substrate of the pdxC (serC) gene product involved in vitamin B6 biosynthesis. *FEBS Letters,* **1996**, *390,* (2), 179-182.
[40] Hibbert, E. G.; Senussi, T.; Costelloe, S. J.; Lei, W.; Smith, M. E.; Ward, J. M.; Hailes, H. C.; Dalby, P. A., Directed evolution of transketolase activity on non-phosphorylated substrates. *Journal of Biotechnology,* **2007**, *131* (4), 425-432.
[41] Ranoux, A.; Karmee, S. K.; Jin, J.; Bhaduri, A.; Caiazzo, A.; Arends, I. W.; Hanefeld, U., Enhancement of the substrate scope of transketolase. *ChemBioChem,* **2012**, *13* (13), 1921-1931.
[42] Tonin, F.; Otten, L. G.; Arends, I. W., NAD+-Dependent Enzymatic Route for the Epimerization of Hydroxysteroids. *ChemSusChem,* **2019**, *12* (13), 3192-3203.
[43] Villegas-Torres, M. F.; Martinez-Torres, R. J.; Cázares-Körner, A.; Hailes, H.; Baganz, F.; Ward, J., Multi-step biocatalytic strategies for chiral amino alcohol synthesis. *Enzyme and Microbial Technology,* **2015**, *81*, 23-30.
[44] Dunbabin, A.; Subrizi, F.; Ward, J. M.; Sheppard, T. D.; Hailes, H. C., Furfurylamines from biomass: transaminase catalysed upgrading of furfurals. *Green Chemistry,* **2017**, *19* (2), 397-404.



7

Thermodynamically and Kinetically Controlled Reactions in Biocatalysis – from Concepts to Perspectives



This chapter is based on

Stefan. R. Marsden, Luuk Mestrom, Duncan G. G. McMillan and Ulf Hanefeld. *ChemCatChem*, **2020**, *12*, 426-437. DOI: 10.1002/cctc201901589.

The enzymatic synthesis of esters and peptides is unfavoured in aqueous solvent systems due to competing hydrolysis. This can be overcome by using energy rich substrate analogues: elimination of a good leaving group temporarily establishes more favourable equilibrium conditions, allowing for (nearly) complete conversion. While kinetically controlled syntheses of esters and peptides in water are common knowledge in biocatalysis textbooks, the prevalence of kinetic control is less well known for other enzyme classes. Here, the general concepts of thermodynamic and kinetic control are illustrated at the example of the synthesis of β-lactam antibiotics and are shown to similarly apply to other enzyme classes. Notably, the enzymatic synthesis of diastereomers shows the same characteristic energy profile as that of Diels-Alder reactions. This allows for the selective synthesis of different diastereomers under either thermodynamically or kinetically controlled conditions. Prospects and pitfalls of this notion are discussed at the example of the thermodynamic epimerization of hydroxysteroids and recent examples of kinetically controlled aldol reactions. Kinetic reaction control can therefore not only be used to increase conversions towards a single product, but also to selectively afford different diastereomers. This review highlights the prevalence of both concepts within the field of biocatalysis.

7.1 Introduction

Catalysis plays a central role for the transition towards more sustainable production processes of chemicals and a circular economy. Biocatalytic approaches are particularly promising due to the typically high chemo- regio- and enantioselectivity of enzymes, which often renders the use of extensive protecting group strategies unnecessary.[1] This feature allows for the design of more direct synthesis routes, effectively decreasing the number of reaction steps that are required to reach the final product.[2] Biocatalytic approaches therefore often allow for substantial improvements in the *E-factor* (which is defined as the amount of waste generated per kilogram of product)[3] by omitting waste intensive purification steps of reaction intermediates in one-pot cascade reactions[4] and avoiding the use of stoichiometric reagents with low atom economy. However, some reactions are inherently limited to low conversions by their thermodynamic equilibrium, which inevitably also leads to a large *E-factor* and threatens the economic viability of the process, even when enzymes are used as catalysts.

Practical solutions to equilibrium limited reactions typically include the addition of one of the substrates in excess, the *in-situ* removal of (side-) products by evaporation or phase separation, or follow up reactions of (side-) products to drive the overall reaction.[5] The enzymatic synthesis of esters or peptides in water is such an example, where equilibrium conditions strongly favour the corresponding alcohol/amine and acid starting materials, for which conversions of less than 1% are commonly observed.[6]

However, energy rich substrate analogues can be used to increase the change in free energy through the elimination of a good leaving group, which then allows for (near) complete conversion of the starting materials under kinetically controlled conditions (Figure 1). In this way, high transient yields can be obtained, but the elimination of an additional leaving group constitutes a drawback from an environmental perspective, as it lowers the overall atom economy.

In the first part of this review, the general concept and characteristics of a kinetically controlled reaction for increased product yields are highlighted at the example of the well-investigated synthesis of β-lactam antibiotics. Based on this, parallels are drawn to other enzyme classes which similarly catalyse kinetically controlled reactions, but have not yet been studied in as much detail. Furthermore, the impact of using substrate analogues on the atom economy is analysed for representative reactions for each enzyme class.

Historically, thermodynamic and kinetic control were first described for Diels-Alder reactions, which can proceed through multiple possible transition states of varying activation energies. Due to the intrinsic energy difference between diastereomers, thermodynamics favour the formation of one particular diastereomer under equilibrium conditions according to

$$K_{eq} = \exp\left(\frac{-\Delta\Delta G}{RT}\right) \qquad (1)$$

(with ΔΔG being the free energy difference between the diastereomers, *R* the universal gas constant, *T* the temperature and K_{eq} the corresponding equilibrium constant). In cases where the desired molecule does not constitute the lowest energy product, its synthesis then essentially needs to be kinetically controlled by close control of the reaction temperature and time.[7]

While thermodynamically and kinetically controlled Diels-Alder type reactions have continued to be the subject of many scientific studies until today,[8] no similarly detailed conceptual studies have been reported for the enzyme catalysed synthesis of diastereomers. Similarities between the energy diagrams of Diels-Alder reactions and the enzymatic synthesis of diastereomers are therefore highlighted in the context of thermodynamic and kinetic control in the second part of our review. Examples from recent literature indicate that while elements of kinetic control were previously observed during the enzymatic synthesis of diastereomers, they rarely are identified as such and therefore remain largely

unexplored. With the continued emergence of thermostable enzymes, biocatalytic transformations are increasingly carried out at elevated temperatures; conditions which are known to favour the formation of thermodynamic products. Insufficient awareness of the possible competition between the target kinetic molecule and the corresponding thermodynamic product could therefore constitute a formidable pitfall during enzyme discovery and evolution for applications in asymmetric synthesis.

7.2 Kinetic control for increased product yields. Concepts and prerequisites

Reactions which only show a small change in Gibbs free energy are intrinsically limited to low conversions, whereas other reactions may also become limited by the process conditions. For instance, water is considered to be an environmentally more benign alternative to organic solvent systems, yet, competing hydrolysis of many target molecules such as esters and peptides can render their synthesis unfavourable under such conditions.[6, 9]

The elimination of a good leaving group from a substrate analogue offers a possible solution, by releasing a larger amount of free energy which temporarily allows for increased conversions (Figure 1). This effect is transient, as the target molecule remains catalytically interconnected with the thermodynamically less favourable reaction. Upon dissipation of the driving force that was released by the elimination of the leaving group, the reverse reaction becomes dominant again and the product concentration ultimately converges back towards the equilibrium concentration of the corresponding initial substrates.

In order to successfully increase conversions by means of kinetically controlled conditions, two separate criteria must therefore be met: Firstly, the substrate analogue must be chosen in a way that elimination of the leaving group releases sufficient additional free energy to increase the equilibrium constant to an extent, which then allows for (near) complete conversion.

Notably, the elimination of a gaseous leaving group (e.g. by decarboxylation) and its release from the reaction mixture can help with driving the reaction, but is not a prerequisite for successful kinetic control. An overview of free energy changes and the corresponding equilibrium constants for representative thermodynamically and kinetically controlled reactions is shown in Table 1 at the end of this section. Secondly, the rate constant for the conversion of the substrate analogue must be sufficiently larger than that of the corresponding reverse reaction, as they are competing throughout the entire course of reaction ($k_1 >> k_2$, Figure 1).

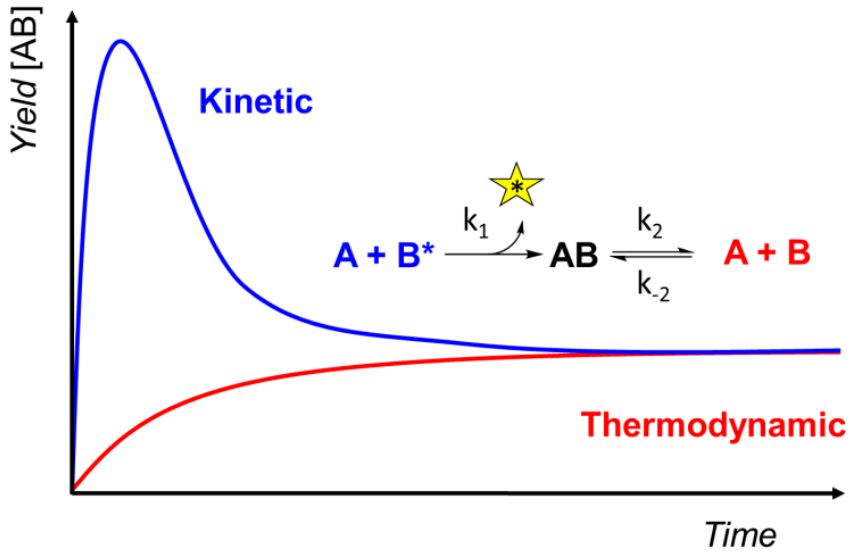
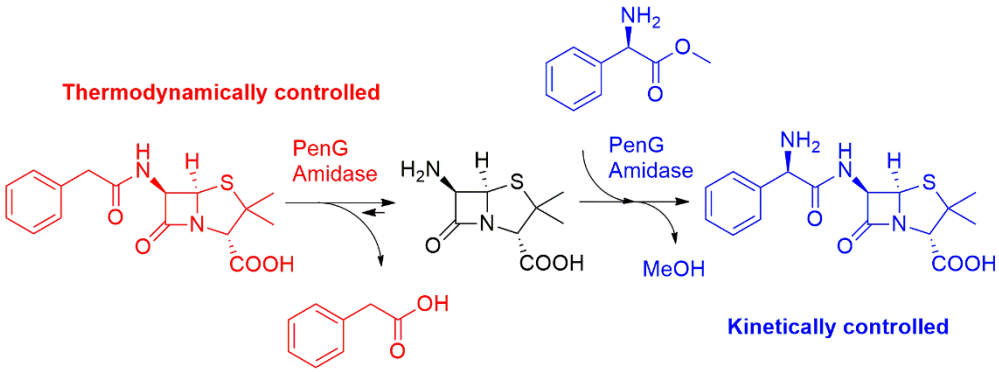


Figure 1: Typical progress curves of product concentrations (AB) in reactions under kinetic control (blue) or thermodynamic control (red). The substrate analogue B* contains a high energy leaving group (*) and its elimination temporarily establishes more favourable equilibrium conditions. Upon dissipation of the initial driving force, the equilibrium between AB and A+B becomes predominant and the product concentration then converges towards that of the corresponding thermodynamically controlled reaction.

The ratio of (k_1/k_2) is an important characteristic of kinetically controlled systems, as it determines the maximum achievable yield.[10] For acyl transferases, the ratio of (k_1/k_2) is termed the synthesis to hydrolysis ratio and is commonly used to compare the efficiency of different enzymes in catalysing a specific kinetically controlled reaction. The rates of conversion for both the substrate analogue (k_1) and the product (k_2) typically follow Michaelis-Menten type kinetics, where k_{cat} and K_M vary for each structural analogue.[6, 9, 11] The individual choice of substrate analogue therefore influences the outcome of the reaction in two ways: first, the release of additional free energy determines the maximum extent of conversion that is possible by *thermodynamics*.[12] Second, its individual rate constant of conversion then determines the ratio of (k_1/k_2), and thereby the maximum transient product yield that is possible by *kinetics*.[9] The ratio of (k_1/k_2) is therefore an important property of the particular catalytic system and should be reported for all kinetically controlled reactions to allow for their comparison. In cases where both rate constants are within the same order of magnitude, addition of several equivalents of the substrate analogue can become necessary to achieve the desired level of conversion.[9] It is worth mentioning that the substrate analogue, which contains the leaving group, is termed the 'donor substrate' in kinetically controlled reactions. This feature must not be confused with its role during the reaction, where it can either act as nucleophile or electrophile. In the following, the application of kinetic control to increase product yields under thermodynamically unfavourable conditions is described for five different enzyme classes to illustrate the widespread utility of this concept within biocatalysis. However, this overview is by no means comprehensive, and many enzymatic reactions involving activated substrates are being continually developed. One such example is the recently described enzymatic Friedel-Crafts acylation with activated acyl donors, which is mechanistically reversible and therefore also is expected to be kinetically controlled.[13]

7.3.1 Amidohydrolases

Enzymatic amide bond formation is one of the key steps in the synthesis of many semi-synthetic β-lactam antibiotics (e.g. amoxicillin or ampicillin) and is being industrially applied annually on a multi-ton scale.[5b] Notably, the production of semi-synthetic β-lactam antibiotics involves both the thermodynamically controlled hydrolysis and subsequent kinetically controlled synthesis of an amide bond, where both reactions are catalysed by the same enzyme (Scheme 1). This reaction requires a high regioselectivity to prevent the undesired hydrolysis of the energy rich β-lactam, which makes enzymes the preferred catalysts for this transformation. Penicillin G (PenG) or penicillin V (PenV) are produced as precursors by hydrolysis of the energy rich β-lactam, which makes enzymes the preferred catalysts for this transformation. Penicillin G (PenG) or penicillin V (PenV) are produced as precursors by fermentation, and hydrolysis of the amide bond affords 6-aminopenicillanic acid (6-APA) as the key building block for β-lactam antibiotics. Conversion of 6-APA with the appropriate acyl donor under kinetically controlled conditions subsequently affords the different members of the β-lactam antibiotic family.[14]



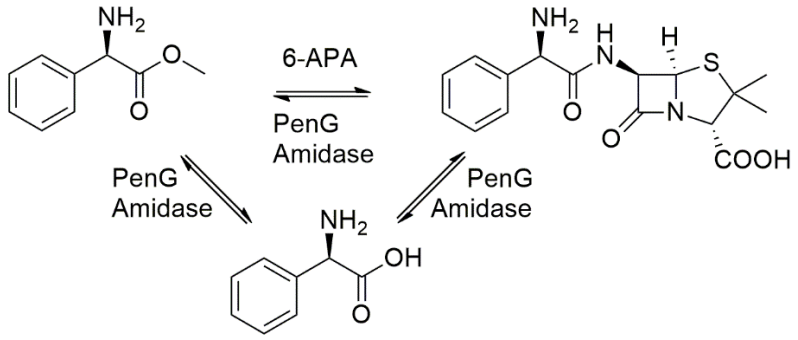
Scheme 1: Thermodynamically controlled hydrolysis of PenG to 6-APA, followed by a kinetically controlled condensation with D-phenylglycine methylester to afford the broadband antibiotic ampicillin.

Due to its industrial relevance, the synthesis of amide bonds is probably the most well studied example of kinetically controlled reactions in biocatalysis, and extensive reviews have previously been published elsewhere.[10, 15]

For thermodynamically controlled reactions, the maximum product yield is determined by the corresponding equilibrium constant and therefore strongly depends on process parameters such as temperature, pH, ionic strength, water activity, presence of organic co-solvents and the molar ratio of substrates used.[10] For instance, elevated reaction temperatures decrease the equilibrium constant for exergonic reactions ($\Delta G < 0$), whereas endergonic reactions ($\Delta G > 0$) become more favourable at higher temperatures according to equation (1). The pH is an important process parameter, as it influences both the change in free energy and the reaction kinetics. In many reactions, the protonation state of functional groups majorly influences their reactivity, and this is also the case for the synthesis of amides. Carboxylate groups have a pK_a in the range of 3, whereas amino groups have a pK_a in the range of 8. Since the reaction occurs between the two uncharged species, a pH value equal to the arithmetic mean of their pK_a values then constitutes a compromise in terms of reactivity. The presence of organic co-solvents can decrease the K_a values of carboxylic acids by up to several orders of magnitude. This improves the relative abundance of reactive substrates by rendering their corresponding pK_a values more similar. [16]

In contrast, the maximum achievable product yield of the corresponding kinetically controlled reaction is determined by the synthesis to hydrolysis ratio (k_1/k_2). This is an intrinsic property of the enzyme, which is derived from its affinity (K_M) and catalytic rate constants (k_{cat}) towards the different substrates that are present in the reaction mixture. In the kinetically controlled synthesis of β-lactam antibiotics, the desired acyl transfer reaction is in constant competition with both the enzyme catalysed and the non-catalytic hydrolysis of the acyl donor and of the product. This increased formation of by-products makes subsequent downstream processing more difficult and constitutes the major drawback to the kinetically

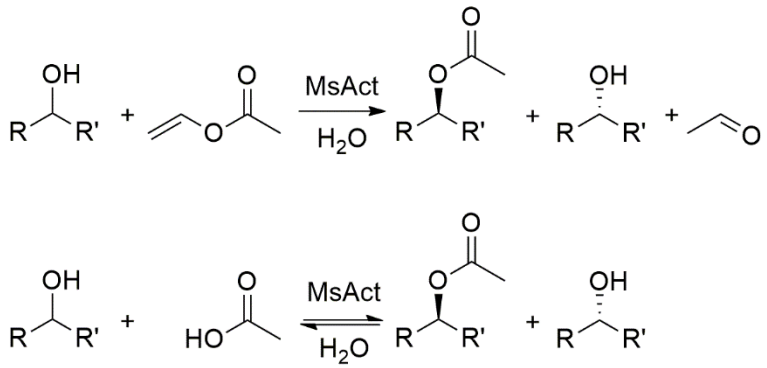
controlled approach (Scheme 2). Hydrolysis of the β-lactam product can be reduced by its extraction into the organic phase in biphasic systems,[17] or by its precipitation.[18] Finally, the synthesis to hydrolysis ratio is also influenced by the method of enzyme immobilisation, where hydrophilic resins seem to give better results than hydrophobic ones.[17, 19]



Scheme 2: Competing hydrolysis of the acyl donor and product to phenylglycine. The formation of salt waste renders downstream processing more difficult and constitutes the major drawback to the kinetically controlled approach.[17]

7.3.2 Acyl transferases

Acyl transferases catalyse the synthesis of esters and amides and are frequently applied in (dynamic) kinetic resolutions for the production of enantiopure amines and alcohols. To render these processes economically feasible, kinetically controlled conditions with an appropriate acyl donor are used to achieve the required levels of conversion. Following the rule of Kazlauskas[20], one of the two enantiomers gets preferentially converted during kinetic resolutions and its separation from the unreacted enantiomer then becomes straightforward. The performance of an enzyme to catalyse kinetic resolution is reflected in the *E ratio*, which can be calculated from the individual rate constants of conversion for each enantiomer.



Scheme 3: Representative acyl transferase catalysed kinetic resolution of a chiral alcohol in aqua using vinyl acetate as an acyl donor. Elimination of vinyl alcohol drives the reaction towards higher transient conversions and tautomerism to acetaldehyde renders the by-product non-nucleophilic. Competing hydrolysis of the product ester requires the reaction to be kinetically controlled (top).[24] The corresponding thermodynamically controlled reaction with acetic acid is limited to low conversions, making it economically unfeasible (bottom).[9]

Notably, the reactivity of the acyl donor must be carefully tuned to match the conditions. On the one hand, a high reactivity of the substrate analogue (k_1) is required to establish an economically acceptable synthesis to hydrolysis ratio (k_1/k_2), whereas an excessively high reactivity results in the non-catalytic conversion of both enantiomers and reduces the overall performance of the kinetic resolution.

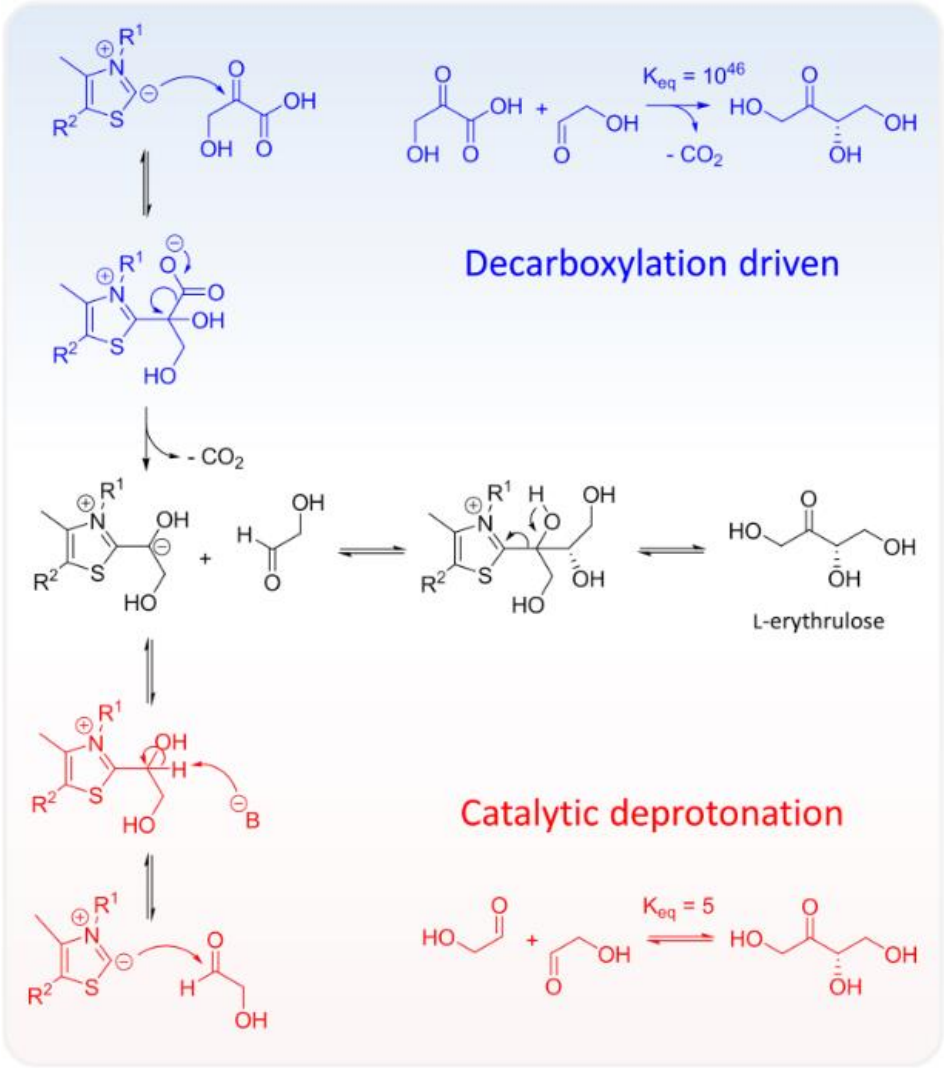
Notably, esters can also be enzymatically synthesised in dry organic solvents or under neat conditions, which offers an alternative solution to the unfavourable thermodynamic equilibrium of ester synthesis *in aqua*.[21] Nonetheless, enzymes generally require a minimum water activity to display their full potential and undesired reverse hydrolysis also remains a challenge in non-aqueous solvent systems.[22] However, enzymatic reactions in dry organic solvents and the influence of water activity lie outside of the scope of this review and more extensive discussions on this topic can be found elsewhere.[21, 23]

7.3.3 Thiamine diphosphate dependent enzymes

Thiamine diphosphate (ThDP) dependent enzymes catalyse the synthesis of chiral α-hydroxyketones (acyloins), which is an important structural motif in many high-value carbohydrates and pharmaceuticals. The reaction comprises of an asymmetric carbon-carbon bond forming step between two aldehyde substrates, of which the donor substrate is rendered a nucleophile following an 'Umpolung' mechanism with the ThDP cofactor. To afford the catalytically active ylide state, the enzyme must stabilise the cofactor in an energetically disfavoured V-conformation that is evolutionarily conserved within this enzyme class.[25] While ThDP dependent enzymes typically convert a large variety of aliphatic-,[26] aromatic-[27] and hydroxyaldehydes[28] as acceptor substrates, they generally have a stricter specificity towards relatively few donor substrates.

The ThDP dependent enzyme transketolase (TK) is a key enzyme of the pentose phosphate pathway, where it catalyses the reversible transfer of a C_2-ketol group from a ketose to an aldose sugar.

For synthetic applications, hydroxypyruvate (HPA) is used as donor substrate, since irreversible decarboxylation directly affords the carbanion in the α,β-dihydroxyethylthiamine diphosphate (DHEThDP) intermediate (Scheme 4, top). In contrast, catalytic deprotonation is required for the formation of the nucleophilic carbanion when glycolaldehyde is being used as the analogous ketol donor without decarboxylation (Scheme 4, bottom). At this stage, both the kinetically and the thermodynamically controlled synthesis of L-erythrulose share the same covalent intermediate DHEThDP, the same mode of C-C bond formation and mechanism of product release.



Scheme 4: Decarboxylation driven synthesis of L-erythrulose from HPA and glycolaldehyde under kinetically controlled conditions (top) and thermodynamically controlled one-substrate reaction with glycolaldehyde (bottom). Both reactions share a common intermediate and mechanism of product release (middle). The product concentration therefore ultimately converges towards the thermodynamic equilibrium of the one-substrate reaction. The pyrimidine and pyrophosphate groups are abbreviated as R^1 and R^2, respectively.

Notably, the reverse splitting of L-erythrulose into two molecules of glycolaldehyde occurs with 100% atom economy. ThDP catalysed reactions therefore constitute a particularly interesting case of kinetic control, where the reverse reaction is an intrinsic property of the product (e.g. the reverse hydrolysis of esters can be circumvented in the absence of water).

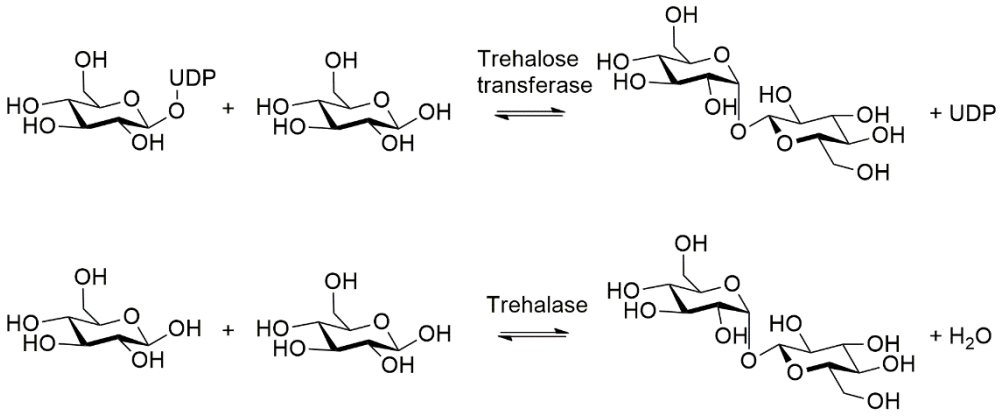
The kinetically controlled synthesis of L-erythrulose displays a high (k_1/k_2) ratio and full thermodynamic equilibration was only observed over extended time periods of several weeks.[26a] The high ratio of (k_1/k_2) previously resulted in inaccurate reports, where product formation was falsely being described as irreversible.[26b, 28b, 29] Based on the reversible splitting of L-erythrulose, an amperometric biosensor has since been developed for the quantitation of ThDP in clinical samples. While ThDP catalysed reactions are essentially irreversible from a practical point of view, this application illustrates the importance of accurately describing similar reactions as being kinetically controlled.[30]

7.3.4 Glycosyl transferases

Glycosyl transferases catalyse the stereoselective synthesis of glycosidic bonds in (poly-) saccharides. In the case of LeLoir glycosyl transferases, activated nucleotide sugars serve as glycosyl donors. The high kinetic stability of glycosidic bonds with respect to non-catalytic hydrolysis makes them a particularly interesting example for the study of thermodynamically and kinetically controlled reactions. The hydrolysis of trehalose into two molecules of glucose is catalysed by the enzyme trehalase. Conversely, trehalose transferase (TreT) does not exhibit any significant hydrolysis activity and therefore displays a near ideal synthesis to hydrolysis ratio with $k_1 >> k_2$.

As described above for acyl transferases, the choice of activated donor substrate plays an important role for the overall performance. For glycosyl transferases, the specific choice of activated nucleotide sugar used for the coupling, such as uridine diphosphate (UDP) or adenosine diphosphate (ADP)

strongly influences the extent of conversion. In contrast to UDP-glucose, the use of ADP-glucose as glycosyl donor releases insufficient free energy during the elimination of ADP to facilitate a shift of the equilibrium constant towards full conversion.[12] Albeit kinetically controlled, the reaction therefore also remains limited by thermodynamics when ADP-glucose is used as glycosyl donor.

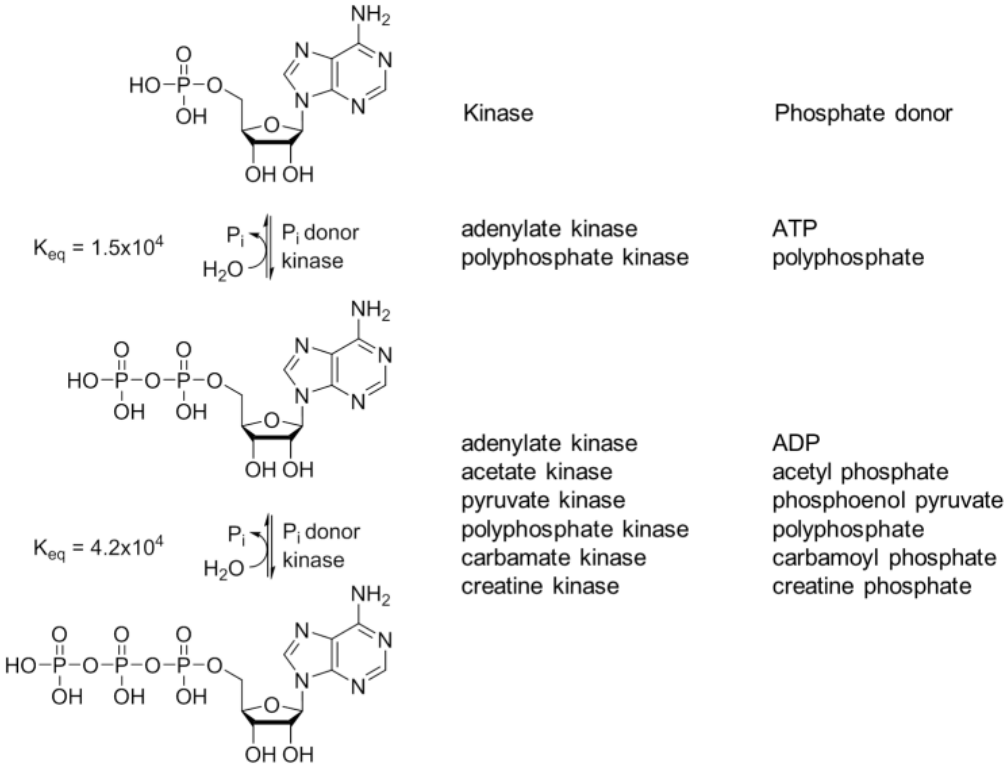


Scheme 5: Trehalose transferase catalysed synthesis of trehalose from a nucleotide sugar (e.g. UDP-glucose) and glucose as acceptor (top) and reverse hydrolysis of trehalose back into the corresponding monomers in aqueous solution (bottom). In nature, the reverse reaction is catalysed by the enzyme trehalase.

7.3.5 Kinases

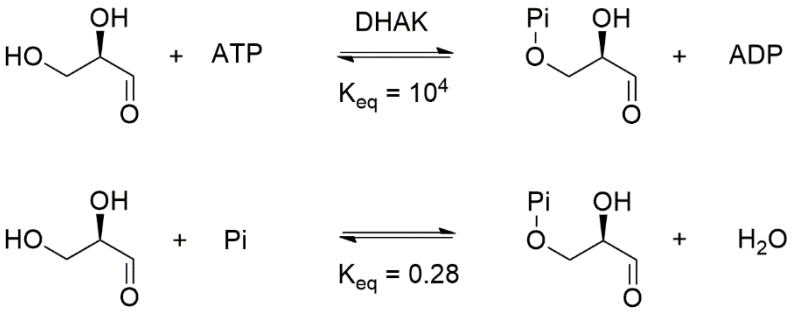
Thermodynamically and kinetically controlled (de-) phosphorylation reactions play a central role in biological systems, where they are used for the activation of biomolecules and in signalling cascades. For example, glycolysis is initiated by the hexokinase catalysed phosphorylation of glucose using ATP. The requirement for phosphorylated substrates during several steps in central metabolism constitutes an elegant regulatory control mechanism.[28b] The kinetically controlled phosphorylation of molecules by kinases functions as an activator, while the thermodynamically preferred hydrolysis of phosphate esters

results in the termination of activity. In the context of biocatalytic conversions, kinases have been extensively used to produce phosphorylated reactants *in situ*.[31] While most kinases use ATP as the preferred phosphate donor, the relatively high price of ATP renders its regeneration from cheaper phosphate sources essential for the economic feasibility of large scale applications. Enzymatic systems for the regeneration of ADP and ATP have therefore extensively been studied.[31, 32] An overview of these systems using various phosphate donors is summarised in Scheme 6.



Scheme 6: Overview of selected kinases and phosphate donors for the successive phosphorylation of AMP via ADP to ATP.[31,32] Equilibrium constants are given for hydrolysis in the reverse direction, which is thermodynamically favoured.

Quantitative ^{31}P-NMR has been demonstrated to be a particularly useful method for the detailed study of the kinetically controlled phosphorylation of D-glyceraldehyde by dihydroxyacetone kinase (DHAK). This phosphorylation reaction was followed over time, allowing for the simultaneous monitoring of substrate conversion, product formation (D-glyceraldehyde phosphate) and its rate of hydrolysis *via* the formation of inorganic phosphate.[33] This reaction is an elegant example of kinetic vs thermodynamic control in that both the regeneration of ATP and its subsequent use in phosphorylation reactions compete with reverse hydrolysis. This infers that kinetic control is a necessity for both steps due to the unfavourable thermodynamic equilibrium of these reactions *in aqua* (Scheme 6, Scheme 7). Given this is one example, the field of kinases clearly warrants further examination in the context of reaction control and limitations.



Scheme 7: Dihydroxyacetone kinase (DHAK) catalysed phosphorylation of D-glyceraldehyde with ATP to D-glyceraldehyde-3-phosphate (G3P) and ADP. Spontaneous hydrolysis of G3P to D-glyceraldehyde and phosphate, with an approximate half-life of 24 hours at pH 6.7 and 298 K.[33]

Table 1: Overview of (transient) equilibrium constants for representative reactions under kinetic control.

Donor	Acceptor	Product	Atom economy	$\Delta rG'0$[a] (kJ/mol)	K_{eq} [b]
Hydroxypyruvate	Glycolaldehyde	Erythrulose	70.6 %	-264.5	$2.2x10^{46}$
Glycolaldehyde	Glycolaldehyde	Erythrulose	100 %	-4.0	5.0
UDP-Glucose	Glucose	Trehalose	45.9 %	-13.2	207
Glucose	Glucose	Trehalose	95 %	11.7	$9.0x10^{-3}$
Vinyl acetate	Benzyl alcohol	Benzyl acetate	77.3 %	-43.2	$3.8x10^{7}$
Acetic acid	Benzyl alcohol	Benzyl acetate	89.3 %	26.6	$2.1x10^{-5}$
ATP	D-glyceraldehyde	D-G3P	28.8 %	-23.3	$1.2x10^{4}$
Phosphate	D-glyceraldehyde	D-G3P	91.5 %	3.1	0.28
D-Phenylglycine methylester	6-APA	Ampicillin	91.6 %	n.a.	n.a.
D-Phenylglycine	6-APA	Ampicillin	95.1 %	n.a.	n.a.

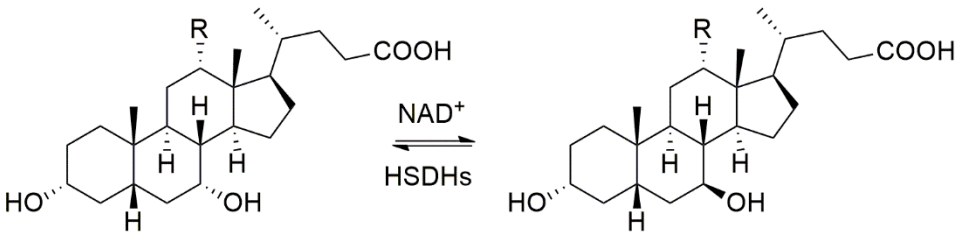
[a] Calculated for pH 7.0 and 0.1 M ionic strength with eQuilibrator.[34] [b] Calculated from the change in Gibb's Free Energy.

7.4 Thermodynamically and kinetically controlled enzymatic syntheses of diastereomers

Enzymes are excellent catalysts that naturally evolved to catalyse chemical conversions with great selectivity and high enantiomeric excess. This is achieved by means of ground state destabilisation, conformational substrate stabilisation, enzyme preorganisation and the stabilisation of transition states through the precise control of spatial arrangements within the active sites.[35] For reactions which start from a prochiral molecule, the two product enantiomers possess the same free energy of formation and the enantiomeric excess remains constant throughout the course of the reaction as an intrinsic property of the active site geometry.

In contrast, diastereomers do not possess the same free energy of formation, which amongst many other properties, allows for their discrimination (e.g. by chromatography or NMR). Most notably, the existence of an energy difference ΔΔG between two diastereomers determines their corresponding equilibrium distribution (i.e. diastereomeric excess) according to equation (1), provided that the reaction conditions allow for their mutual interconversion.

This concept has been elegantly applied in the thermodynamic epimerization of hydroxysteroids for the production of ursodeoxycholic acid, an active pharmaceutical ingredient that is being used for the treatment of cholestatic diseases.[36] A combination of two stereocomplementary hydroxysteroid dehydrogenases (7α- and 7β HSDHs) was used to chemically interconnect both diastereomers *via* the corresponding ketone in a redox neutral fashion, and conversions in excess of 90% towards the target diastereomer were achieved with their energy difference acting as the sole driving force.[37] However, in cases where the target diastereomer does not constitute the thermodynamic product, its synthesis then effectively needs to be kinetically controlled to obtain a high diastereomeric excess.



R = H (CDCA); R = OH (CA)

Scheme 8: Thermodynamic epimerization of the 7-hydroxyl group in cholic (CA) or chenodeoxycholic acid (CDCA) by application of two stereocomplementary hydroxysteroid dehydrogenases (7α- and 7β-HSDHs) in a redox neutral fashion. An energy difference of 1.23 kJ/mol under standard conditions was calculated for the two diastereomers (K_{eq}=1.64), which acts as the sole driving force for epimerization.[38]

The thermodynamically and kinetically controlled synthesis of diastereomers has extensively been studied for Diels-Alder reactions since their discovery in 1929,[7] and has remained the topic of extensive studies until today.[8] In this type of reaction, the conversion selectively proceeds through the lowest energy transition state at low temperatures, whereas higher temperatures allow for it to proceed through both transition states and therefore to equilibrate towards the formation of the thermodynamic product (Figure 2).

In the following, parallels are drawn between the energy diagram of Diels-Alder reactions (Figure 2) and that for the enzymatic synthesis of diastereomers (Figure 3). For simplicity reasons, the introduction of one new chiral centre to afford two diastereomers will be discussed at the example of a hypothetical reaction, that is catalysed by a single enzyme and affords the *(S)* configuration at the newly formed stereocentre as the thermodynamic product. In spite of their generally high stereoselectivity, enzymes are not perfect catalysts and therefore, to some extent, catalyse the formation of both stereoisomers. Since most enzymatic reactions are mechanistically reversible, the product diastereomers are rendered dynamically interconnected (Figure 3).

The equilibration towards the thermodynamic product was previously already demonstrated for the mechanistically interconnected, kinetically controlled synthesis of L-erythrulose using the enzyme transketolase.[26a]

Due to the asymmetry of active sites, the pro-*R* and pro-*S* transition states are stabilised by different chemical environments and thereby also display different activation energies (E_a).[35] By comparison, the enzymatic synthesis of diastereomers shows the same characteristic energy diagram as Diels-Alder reactions (Figure 2, Figure 3).

In cases where the *(S)* configured diastereomer constitutes the thermodynamic product, an *(R)*-selective enzyme will initially afford the *(R)* configured product under kinetically controlled conditions.

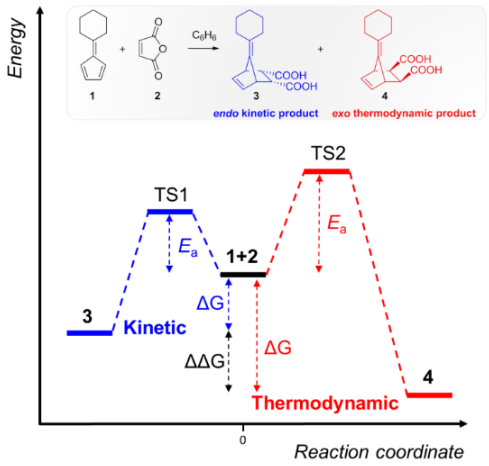


Figure 2: Qualitative energy diagram for the Diels-Alder reaction of fulvene **1** and maleic anhydride **2** to selectively afford *endo* **3** or *exo* **4** configured products under either thermodynamically or kinetically controlled conditions.

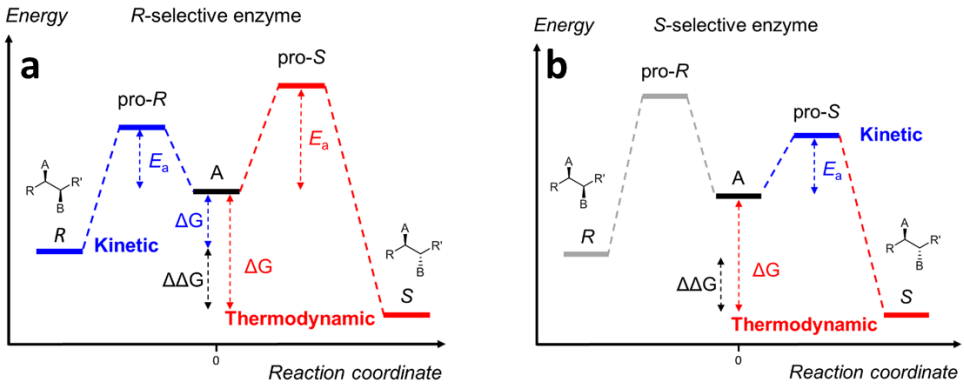
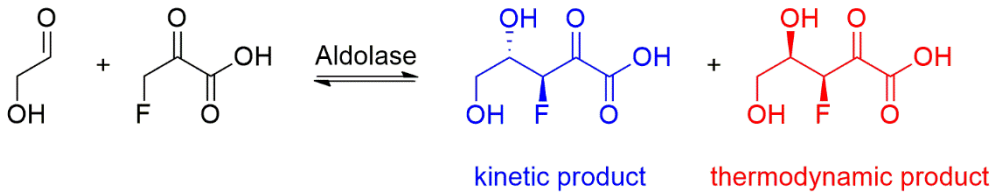


Figure 3: Qualitative energy diagrams for a hypothetical reaction, in which the (*S*)-configured diastereomer is lowest in energy. **(a)** An *(R)*-selective enzyme affords the *(R)*-configuration as the kinetic product, with subsequent equilibration towards the *(S)*-configuration under thermodynamically controlled conditions. **(b)** An *(S)*-selective enzyme (bottom), the *(S)* diastereomer is directly formed as the kinetic product. Under thermodynamically controlled conditions, the final diastereomeric excess is identical for both enzymes.

If the reaction is allowed to proceed, subsequent epimerization into the *(S)* diastereomer will occur and establish the corresponding thermodynamic equilibrium distribution of diastereomers based on their difference in free energy. Similarly, an *(S)* selective enzyme will initially afford the *(S)* configured product under kinetically controlled conditions with high stereoselectivity as a reflection of the enzyme's active site geometry. Nonetheless, prolonged reaction times will establish the same distribution of diastereomers for both an *(R)* or *(S)* selective enzyme under thermodynamic conditions. Thus, the diastereomeric excess will also decrease over time for an *(S)* selective enzyme.

For example, the conversion of fluoropyruvate by a variety of type I and type II aldolases initially affords a single product with high diastereomeric excess, but this value was observed to decrease over time and the reactions needed to be kinetically controlled (Scheme 9).[38] In another study, lower enzyme loadings increased the diastereomeric excess under otherwise identical conditions at the expense of conversion.[39] Thus, the absolute configuration at the newly formed chiral centre *(S)* does not remain constant and changes over time towards that of the thermodynamic product.



Scheme 9: Aldol coupling of fluoropyruvate and glycolaldehyde, catalysed by three different aldolases (*Ec*GarL, *Ec*RhmA, *Ec*HpcH). The anti-configured product is obtained with high diastereomeric excess (71–93%) at low catalyst loadings (0.01 mol%), while only a low to moderate diastereomeric excess (39–59%) is obtained with a higher catalyst loading (0.05 mol%).[39]

A qualitative progression curve of the absolute configuration for the previously discussed hypothetical reaction is shown in Figure 4. Both the position of the final equilibrium and the rate of progression towards it strongly depend on the specific reaction parameters (temperature, pH, enzyme loading). Single measurements of the diastereomeric excess therefore only allow to a limited extent for conclusions regarding the enzyme's stereoselectivity. This could even lead to a falsely assigned stereoselectivity for the enzyme, as the actual position within this progress curve remains unknown.

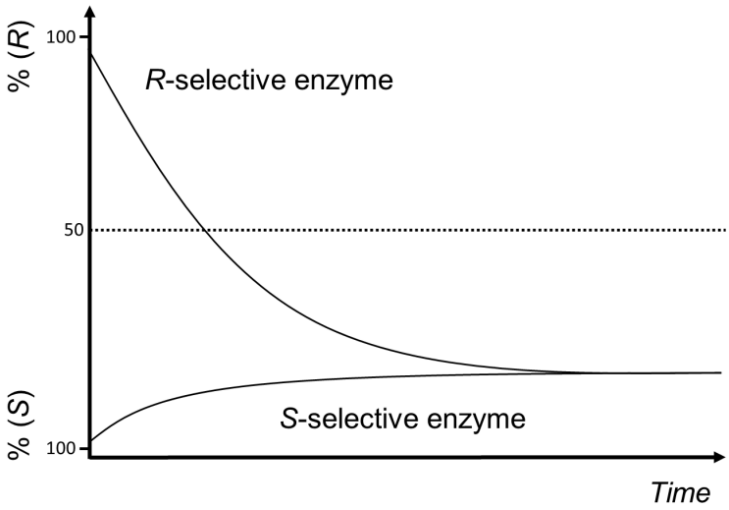


Figure 4: Progression of the absolute configuration at the newly formed chiral centre over time for the hypothetical reaction discussed in Figure 3.

Formerly, biocatalytic conversions were typically carried out within a narrow temperature range close to physiological conditions, due to the frequent temperature sensitivity of microorganisms and wild-type enzymes. These conditions generally favour the synthesis of kinetic products. With the emergence of computational algorithms to enhance the thermostability of enzymes,[40] and the plethora of genomes that are currently available from thermophilic organisms,

biocatalytic conversions are becoming increasingly compatible with elevated reaction temperatures.[41] As a consequence, reactions are increasingly at risk to also proceed through higher energy transition states, which circumvents kinetic control and more rapidly establishes the thermodynamic equilibrium distribution of products.

While it is easy to conceptualize transition states as well-defined structures, *transition state landscapes* better capture their dynamic and structurally diverse nature.[42] A detailed QM/MM analysis of enzymes belonging to the alkaline phosphatase superfamily described the stabilisation of multiple loose transition states within single active sites.[43] This study found that alkaline phosphatases are able to recognise and stabilise multiple transition states without undergoing larger structural rearrangements. This suggests that elevated reaction temperatures could similarly also impact an enzyme's regioselectivity by allowing conversions to proceed through more distant transition states that are higher in energy. Elevated reaction temperatures could therefore potentially impact the outcome of many reactions, ranging from a switch in absolute configuration, to the formation of entirely new products as a result of a changed regioselectivity.

7.5 Conclusions

Using the elimination of a good leaving group in substrate analogues as driving force, kinetically controlled conditions enable enzyme catalysed conversions which otherwise would not be economically viable under thermodynamically controlled conditions. While the use of substrate analogues comes at the cost of a lower atom economy, this is generally outweighed by the benefits that come with increased levels of conversion and improved chemo-, regio- and enantioselectivity. In this way, kinetically controlled reactions have found widespread applications for the synthesis of chiral building blocks and high value chemicals on large scale.

The enzyme catalysed synthesis of diastereomers shows the same characteristic energy diagram as that for Diels-Alder reactions and enzymatic conversions similarly proceed initially through the lowest energy transition state at low temperatures.

However, enzymes are not perfect catalysts and therefore mechanistically interconnect the synthesis of both diastereomers. Ultimately, enzymes establish the equilibrium distribution of diastereomers that is intrinsically determined by their difference in free energy. While this may be desirable in cases like the thermodynamic epimerization of hydroxysteroids, the target molecule does not necessarily always constitute the thermodynamic product, and its synthesis then needs to be kinetically controlled. While the absolute configuration of the newly formed chiral centre *(S)* initially is determined by the enzymes stereoselectivity under kinetically controlled conditions, thermodynamic equilibration can lead to the inversion of the stereocentre. Single measurements of the diastereomeric excess therefore only provide limited information and could even lead to an enzyme's stereoselectivity being falsely assigned. This notion is of particular relevance, as biocatalytic conversions are increasingly conducted at elevated temperatures; conditions which are known to favour the formation of thermodynamic products.

Acknowledgements

The authors would like to thank all group members, past and present that have participated in and contributed to our research effort. L. M. is grateful for the generous sponsoring via the ERA-IB scheme, grant ERA-IB-15-110.

References

[1] K. Faber, *Biotransformations in organic chemistry*, 7 ed., Springer, **2018**.

[2] L. Aitao, S. Au, S. Bartsch, D. Beecher, A. Boffi, *Science of Synthesis: Biocatalysis in Organic Synthesis, Vol. 2*, Georg Thieme Verlag, **2015**.

[3] a) R. A. Sheldon, *C. R. Acad. Sci.* **2000**, *3*, 541-551; b) R. A. Sheldon, *Green Chem.* **2007**, *9*, 1273-1283; c) R. A. Sheldon, *Green Chem.* **2017**, *19*, 18-43.

[4] W.-D. Fessner, *New Biotechnol.* **2015**, *32*, 658-664.

[5] a) A. Liese, K. Seelbach, C. Wandrey, *Industrial biotransformations*, John Wiley & Sons, **2006**; b) K. Buchholz, V. Kasche, U. T. Bornscheuer, *Biocatalysts and Enzyme Technology*, John Wiley & Sons, **2012**.

[6] M. Subileau, A. Jan, J. Drone, C. Rutyna, V. Perrier, E. Dubreucq, *Catal. Sci. Technol.* **2017**, *7*, 2566-2578.

[7] a) R. Woodward, H. Baer, *J. Am. Chem. Soc.* **1944**, *66*, 645-649; b) O. Diels, K. Alder, *Ber. Dtsch. Chem. Ges. B* **1929**, *62*, 2081-2087.

[8] a) K. K. Borisova, E. A. Kvyatkovskaya, E. V. Nikitina, R. R. Aysin, R. A. Novikov, F. I. Zubkov, *J. Org Chem.* **2018**, *83*, 4840-4850; b) K. K. Borisova, E. V. Nikitina, R. A. Novikov, V. N. Khrustalev, P. V. Dorovatovskii, Y. V. Zubavichus, M. L. Kuznetsov, V. P. Zaytsev, A. V. Varlamov, F. I. Zubkov, *Chem. Commun.* **2018**, *54*, 2850-2853; c) D. L. Boger, S. M. Weinreb, *Hetero Diels-Alder methodology in organic synthesis, Vol. 47*, Elsevier, **2012**; d) H. B. Kagan, O. Riant, *Chem. Rev.* **1992**, *92*, 1007-1019; e) K. C. Nicolaou, S. A. Snyder, T. Montagnon, G. Vassilikogiannakis, *Angew. Chem. Int. Ed.* **2002**, *41*, 1668-1698; f) G. Brieger, J. N. Bennett, *Chem. Rev.* **1980**, *80*, 63-97.

[9] L. Mestrom, J. G. Claessen, U. Hanefeld, *ChemCatChem* **2019**, *11*, 2004-2010.

[10] V. Kasche, *Enzyme Microb. Technol.* **1986**, *8*, 4-16.

[11] a) A. H. Jan, É. Dubreucq, M. Subileau, *ChemBioChem* **2017**, *18*, 941-950; b) A. H. Jan Deniau, M. Subileau, E. Dubreucq, *ChemBioChem* **2018**, *19*, 1839-1844.

[12] a) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Appl. Environ. Microbiol.* **2019**, *85*, e03084-03018; b) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Appl. Environ. Microbiol.* **2019**, *85*, e00942-00919; c) S.-I. Ryu, J.-E. Kim, E.-J. Kim, S.-K. Chung, S.-B. Lee, *Process Biochem.* **2011**, *46*, 128-134.

[13] N. G. Schmidt, T. Pavkov-Keller, N. Richter, B. Wiltschi, K. Gruber, W. Kroutil, *Angew. Chem. Int. Ed.* **2017**, *56*, 7615-7619.

[14] M. van der Mey, E. de Vroom, *Biorg. Med. Chem. Lett.* **1994**, *4*, 345-348.

[15] a) A. Bruggink, E. C. Roos, E. de Vroom, *Org. Process Res. Dev.* **1998**, *2*, 128-133; b) A. Illanes, P. Valencia, in *Current Developments in Biotechnology and Bioengineering*, Elsevier, **2017**, pp. 267-305.

[16] a) R. Fernández-Lafuente, G. Alvaro, R. M. Blanco, J. M. Guisán, *Appl. Biochem. Biotechnol.* **1991**, *27*, 277-290;

b) R. Fernández-Lafuente, C. M. Rosell, J. M. Guisán, *Enzyme Microb. Technol.* **1991**, *13*, 898-905.

[17] O. Hernández-Jústiz, M. Terreni, G. Pagani, J. Garcıa, J. Guisán, R. Fernández-Lafuente, *Enzyme Microb. Technol.* **1999**, *25*, 336-343.

[18] R. Ulijn, L. De Martin, P. Halling, B. Moore, A. Janssen, *J. Biotechnol.* **2002**, *99*, 215-222.

[19] a) M. Terreni, D. Ubiali, T. Bavaro, M. Pregnolato, R. Fernández-Lafuente, J. M. Guisán, *Appl. Microbiol. Biotechnol.* **2007**, *77*, 579-587; b) D. A. Cecchini, R. Pavesi, S. Sanna, S. Daly, R. Xaiz, M. Pregnolato, M. Terreni, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 1491-1500; c) M. Terreni, G. Pagani, D. Ubiali, R. Fernandez-Lafuente, C. Mateo, J. M. Guisán, *Biorg. Med. Chem. Lett.* **2001**, *11*, 2429-2432; d) L. R. Gonçalves, R. Sousa Jr, R. Fernandez-Lafuente, J. M. Guisan, R. L. Giordano, R. C. Giordano, *Biotechnol. Bioeng.* **2002**, *80*, 622-631; e) O. Justiz, R. Fernandez-Lafuente, J. Guisan, P. Negri, G. Pagani, M. Pregnolato, M. Terreni, *J. Org. Chem.* **1997**, *62*, 9099-9106.

[20] R. J. Kazlauskas, A. N. Weissfloch, A. T. Rappaport, L. A. Cuccia, *J. Org. Chem.* **1991**, *56*, 2656-2665.

[21] a) P. Hara, U. Hanefeld, L. T. Kanerva, *Green Chem.* **2009**, *11*, 250-256; b) P. Hara, U. Hanefeld, L. T. Kanerva, *J. Mol. Catal. B: Enzym.* **2008**, *50*, 80-86.

[22] I. Svensson, E. Wehtje, P. Adlercreutz, B. Mattiasson, *Biotechnol. Bioeng.* **1994**, *44*, 549-556.

[23] A. L. Margolin, A. M. Klibanov, *J. Am. Chem. Soc.* **1987**, *109*, 3802-3804.

[24] N. de Leeuw, G. Torrelo, C. Bisterfeld, V. Resch, L. Mestrom, E. Straulino, L. van der Weel, U. Hanefeld, *Adv. Synth. Catal.* **2018**, *360*, 242-249.

[25] a) W. Shin, J. Pletcher, G. Blank, M. Sax, *J. Am. Chem. Soc.* **1977**, *99*, 3491-3499; b) A. Schellenberger, Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. **1998**, *1385*, 177-186; c) P. Asztalos, C. Parthier, R. Golbik, M. Kleinschmidt, G. Hübner, M. S. Weiss, R. Friedemann, G. Wille, K. Tittmann, *Biochemistry* **2007**, *46*, 12037-12052.

[26] a) S. R. Marsden, L. Gjonaj, S. J. Eustace, U. Hanefeld, *ChemCatChem* **2017**, *9*, 1808-1814; b) A. Cázares, J. L. Galman, L. G. Crago, M. E. Smith, J. Strafford, L. Ríos-Solís, G. J. Lye, P. A. Dalby, H. C. Hailes, *Org. Biomol. Chem.* **2010**, *8*, 1301-1309; c) D. Yi, T. Saravanan, T. Devamani, F. Charmantray, L. Hecquet, W.-D. Fessner, *Chem. Commun.* **2015**, *51*, 480-483.

[27] T. Saravanan, M.-L. Reif, D. Yi, M. Lorillière, F. Charmantray, L. Hecquet, W.-D. Fessner, *Green Chem.* **2017**, *19*, 481-489.

[28] a) E. G. Hibbert, T. Senussi, S. J. Costelloe, W. Lei, M. E. Smith, J. M. Ward, H. C. Hailes, P. A. Dalby, *J. Biotechnol.* **2007**, *131*, 425-432; b) A. Ranoux, S. K. Karmee, J. Jin, A. Bhaduri, A. Caiazzo, I. W. Arends, U. Hanefeld, *ChemBioChem* **2012**, *13*, 1921-1931; c) M. Lorillière, R. Dumoulin, M. L'enfant, A. Rambourdin, V. Thery, L. Nauton, W.-D. Fessner, F. Charmantray, L. Hecquet, *ACS Catalysis* **2019**.

[29] a) G. A. Kochetov, O. N. Solovjeva, Biochim. Biophys. Acta, Proteins Proteomics **2014**, *1844*, 1608-1618; b) J. Abdoul Zabar, M. Lorillière, D. Yi, T. Saravanan, T. Devamani, L. Nauton, F. Charmantray, V. Hélaine, W. D. Fessner, L. Hecquet, *Adv. Synth. Catal.* **2015**, *357*, 1715-1720; c) Y. Kobori, D. C. Myles, G. M. Whitesides, *J. Org. Chem.* **1992**, *57*, 5899-5907; d) K. Benaissi, V. Hélaine, V. Prévot, C. Forano, L. Hecquet, *Adv. Synth. Catal.* **2011**, *353*, 1497-1509; e) J. Sukumaran, U. Hanefeld, *Chem. Soc. Rev.* **2005**, *34*, 530-542.

[30] M. Halma, B. Doumèche, L. Hecquet, V. Prévot, C. Mousty, F. Charmantray, *Biosensors Bioelectron.* **2017**, *87*, 850-857.

[31] J. N. Andexer, M. Richter, *ChemBioChem* **2015**, *16*, 380-386.

[32] a) H. K. Chenault, E. S. Simon, G. M. Whitesides, *Biotechnol. Genet. Eng. Rev.* **1988**, *6*, 221-270; b) H. Zhao, W. A. Van Der Donk, *Curr. Opin. Biotechnol.* **2003**, *14*, 583-589.

[33] D. Gauss, I. Sánchez-Moreno, I. Oroz-Guinea, E. García-Junceda, R. Wohlgemuth, *Eur. J. Org. Chem.* **2018**, *2018*, 2892-2895.

[34] E. Noor, H. S. Haraldsdóttir, R. Milo, R. M. Fleming, *PLoS Comp. Biol.* **2013**, *9*, e1003098.

[35] B. Vögeli, T. J. Erb, *Curr. Opin. Chem. Biol.* **2018**, *47*, 94-100.

[36] a) T. Ikegami, Y. Matsuzaki, *Hepatol. Res.* **2008**, *38*, 123-131; b) F. Tonin, I. W. Arends, *Beilstein J. Org. Chem.* **2018**, *14*, 470-483.

[37] F. Tonin, L. G. Otten, I. W. Arends, *ChemSusChem* **2019**, *12*, 3192-3203.

[38] a) J. Stockwell, A. D. Daniels, C. L. Windle, T. A. Harman, T. Woodhall, T. Lebl, C. H. Trinh, K. Mulholland, A. R. Pearson, A. Berry, *Org. Biomol. Chem.* **2016**, *14*, 105-112; b) C. L. Windle, A. Berry, A. Nelson, *Curr. Opin. Chem. Biol.* **2017**, *37*, 33-38.

[39] J. Fang, D. Hait, M. Head-Gordon, M. C. Y. Chang, *Angew. Chem. Int. Ed.* **2019**, *58*, 11841-11845.

[40] M. Musil, J. Stourac, J. Bendl, J. Brezovsky, Z. Prokop, J. Zendulka, T. Martinek, D. Bednar, J. Damborsky, *Nucleic Acids Res.* **2017**, *45*, W393-W399.

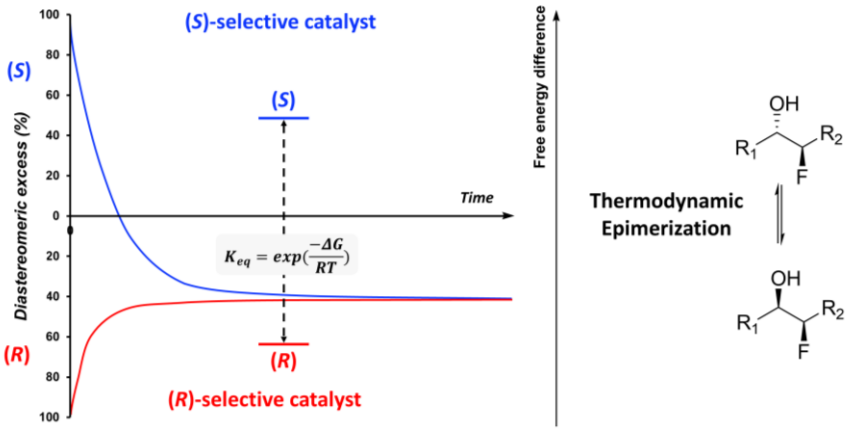
[41] G. Haki, S. Rakshit, *Bioresour. Technol.* **2003**, *89*, 17-34.

[42] a) B. Ma, S. Kumar, C.-J. Tsai, Z. Hu, R. Nussinov, *J. Theor. Biol.* **2000**, *203*, 383-397; b) D. G. Truhlar, *Arch. Biochem. Biophys.* **2015**, *582*, 10-17.

[43] G. Hou, Q. Cui, *J. Am. Chem. Soc.* **2013**, *135*, 10457-10469.

8

Thermodynamics Determine the Diastereochemical Outcome of Catalytic Reactions



This chapter is based on

Stefan R. Marsden, Luuk Mestrom, Hein J. Wijma, Sander J. Noordam, Duncan G. G. McMillan and Ulf Hanefeld.

ChemCatChem, **2021**. DOI: 10.1002/cctc.202100178.

Diastereomers are characterized by an intrinsic energy difference, and thermodynamics dictate their distribution within a dynamic equilibrium. The characteristic mechanistic reversibility and non-ideal stereoselectivity of catalysts therefore simultaneously promote both synthesis and epimerization of products during the formation of diastereomers. This feature can even result in the thermodynamic inversion of a chiral centre against the catalyst's stereoselectivity. Here, we provide a comprehensive experimental and theoretical study of factors that govern thermodynamic epimerization in catalysis, using enzymes as example. Our analysis highlights, that the deduction of a catalyst's stereoselectivity based on the absolute configuration of the isolated product constitutes a potential pitfall. The selective formation of either the thermodynamic-, or the kinetic product is less determined by the catalyst, but rather by the reaction conditions. Next to low temperatures, a high maximal extent of conversion was identified to promote kinetically controlled conditions. For bimolecular reactions, conversions can be conveniently modulated *via* the use of one substrate in excess. Quantum mechanical calculations accurately predicted the diastereomeric excess under equilibrium conditions, which opens the prospect of a rational choice between thermodynamic and kinetic reaction control at an early stage of process design. Our findings are of critical importance for multi-step syntheses of stereocomplex molecules *via* catalytic cascade reactions or artificial metabolic pathways, as the final stereochemistry may be determined by the absolute configuration of the product that is overall lowest in energy.

8.1 Introduction

Biological systems are inherently chiral due to the unique structures of DNA, proteins and metabolites, and two enantiomers of the same compound often display different biological properties.[1] This notion renders the development of catalysts that display a high chemo-, regio- and enantioselectivity fundamentally important for applications within life sciences. Kinetic reaction control is a widely prevalent approach in organic synthesis to modulate the product selectivity of a reaction. It is conceptually based on the faster formation of a reaction intermediate (i.e. the kinetic product), over formation of the lowest energy (i.e. thermodynamic) product.[2] Continuous monitoring of the reaction's progress is essential to circumvent equilibration toward the most stable product under thermodynamically controlled conditions.[3] For instance, the stereoselective synthesis of enantiomers is kinetically controlled, while the racemic mixture constitutes the thermodynamic product.[4] Enantiomers possess the same free energy of formation ($\Delta G = 0$), and an increase in entropy promotes their racemization ($K_{eq} = 1$, *ee*% = 0) according to

$$\Delta G = -RT\ln K_{eq}.^{[5]} \quad (1)$$

Conversely, diastereomers are characterized by an intrinsic energy difference; this feature determines their distribution within a dynamic equilibrium. This is for instance the case with readily interconverting conformational isomers, and their irreversible conversion following Curtin-Hammet/Winstein-Holness (CH/WH) type kinetics has been extensively investigated since the 1950s.[6, 7] The CH/WH principle states, that for k_{AC},k_{BD} >> k_{AB},k_{BA}, the product distribution of C,D reflects the initial distribution of diastereomeric conformers A,B (Figure 1a). The irreversible nature of reactions that are described by CH/WH type kinetics renders these reactions kinetically controlled and prevents their subsequent thermodynamic equilibration. For the reverse case k_{AC},k_{BD} << k_{AB},k_{BA}, the distribution of CD is also constant and determined by k_{AC},k_{BD}. In contrast, mechanistically reversible, non-catalytic Diels-Alder reactions afford the product diastereomers within a dynamic equilibrium.[8] This allows for the selective isolation

of either the thermodynamic or the kinetic reaction product *via* modulation of the process conditions.[9] At low temperatures, Diels-Alder reactions selectively proceed through the lowest energy transition state to afford the kinetic product, while prolonged reaction times and elevated temperatures result in the formation of the thermodynamic product.[8, 9] The diastereomeric excess (*de*) is then based on the corresponding energy difference between the diastereomers.

Analogous to non-catalytic Diels-Alder reactions, the diastereoselectivity of a catalyst can be translated into a similar energy diagram (Figure 1b,c).[3] For simplicity reasons, the second stereocentre (X), that is required for a molecule to be a diastereomer, is considered to be fixed as either (*R*)- or (*S*)-configured. However, the discussed concepts similarly apply to catalysts that display no enantioselectivity, and therefore afford (X) in both configurations. In this interpretation, the activation energies E_a of the corresponding pro-(*R*) and pro-(*S*) transition states reflect the catalyst's diastereoselectivity.[10] The characteristic mechanistic reversibility and non-ideal stereoselectivity of catalysts then establish a dynamic equilibrium between the different product diastereomers. Catalysts therefore simultaneously promote both the synthesis and epimerization of products during the formation of diastereomers. This is conceptually similar to Diels-Alder reactions, and opposite to irreversible, CH/WH type systems. Under thermodynamic conditions, the stereochemical outcome for the catalytic synthesis of diastereomers is therefore independent of the catalyst, and may be opposite to its stereoselectivity.

Declaredly, epimerization effects were previously reported at the specific example of threonine aldolases.[11-18] Temperature and ratio of substrates were described to have a strong influence on the outcome of the reaction, indicating the influence of thermodynamics on the diastereoselectivity. Yet, surprisingly, similar studies with other enzymes and a broader discussion of thermodynamic epimerization effects in catalysis, such as their implications for multi-step cascade reactions, remain largely absent in literature.

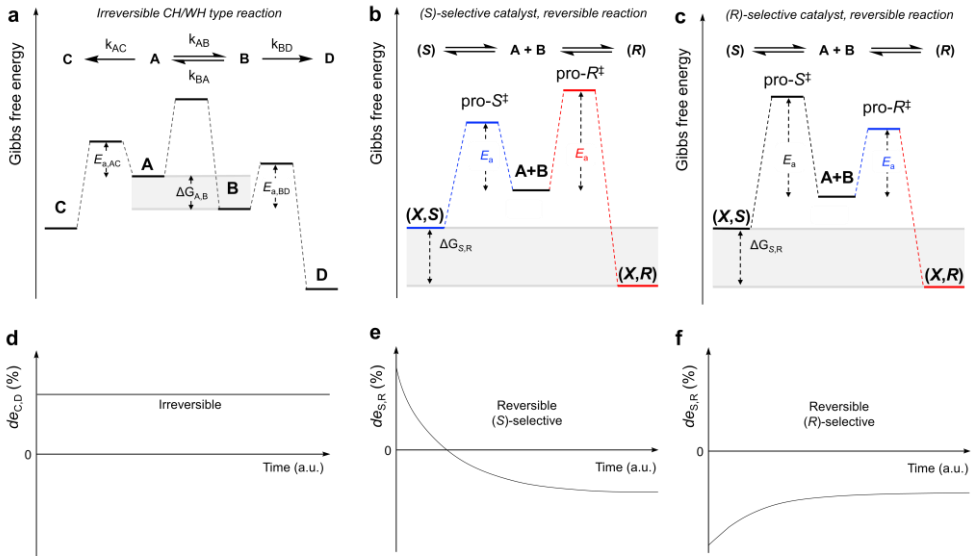


Figure 1: Qualitative energy diagrams and related diastereomeric excess time courses. **(a,d)** Curtin-Hammett type irreversible conversion of interconverting conformers AB. For k_{AB},k_{BA} << k_{BD},k_{AC}, the distribution of CD is constant and reflects the initial distribution of A and B. **(b,c)** Hypothetical reaction, in which the (*X,R*)-configured diastereomer constitutes the thermodynamic product, catalysed by either (*X,R*)- or (*X,S*)-selective catalysts. The second stereocentre (X, for diastereomers required) is fixed and can be either (*R*)- or (*S*)-configured. The catalyst's stereoselectivity is reflected by a lower activation energy E_a. **(e,f)** The kinetic product is initially afforded according to the catalyst's diastereoselectivity. Simultaneous epimerization ultimately establishes the same thermodynamic distribution of diastereomers according to $\Delta G_{R,S}$. For (*S*)-selective catalysts, this inverts the stereochemical outcome **(e)**. For (*R*)-selective catalysts, it reduces the diastereomeric excess **(f)**.

Here, we present a theoretical and experimental analysis of general factors that influence the extent of thermodynamic epimerization, such as temperature, catalyst stereoselectivity, free energy difference between product diastereomers and the particular stoichiometry of substrates. Quantum mechanical methods were

applied to predict the identity and associated diastereomeric excess of the thermodynamic product under equilibrium conditions. Finally, catalytic single and two-step conversions were used as model systems to study the effect of thermodynamic epimerization for multi-step cascade reactions. In this manner, we aim to provide a comprehensive overview of thermodynamic epimerization effects for catalysis.

8.2 Results and Discussion

Time course studies

The stereocomplementary hydroxyketoacid aldolases from *Sphingomonas wittichii* (*Sw*HKA) and *Burkholderia phytofirmans* (*Bp*HKA) selectively catalyse an aldol reaction between fluoropyruvate as the donor substrate, and a wide variety of aldehydes as acceptor substrates.[19] This conversion creates two chiral centres in one step from achiral substrates, circumventing possible interferences from a (dynamic) kinetic resolution of the substrates. This makes it an ideal system for monitoring thermodynamic epimerization in a time-related process (Figure 2).

The extent of conversion and absolute configurations were both quantitatively monitored over time by ^{19}F-NMR. At low levels of conversion, spectra in which the signal-to-noise ratio of the minor product peak was below 2:1 were discarded, in order to reliably determine the diastereomeric excess in all cases. Stereoselective reduction of the ketoacid aldol products with L-lactic dehydrogenase was applied to differentiate the enantiomers by their conversion into unique diastereomers. This allowed to determine both the enantiomeric excess, and to assign the absolute configuration for the fluorine moiety *via* its $^{3}J_{HF}$ coupling constant to the newly formed stereocentre.

In agreement with previous reports,[19, 20] both *Sw*HKA and *Bp*HKA were found to be highly (*3S*)-selective, and the minor (*3R*)-configured enantiomers were below the limits of reliable detection under the conditions examined. However, both

enzymes displayed a moderate diastereoselectivity with respect to the configuration of the C-4 hydroxyl group. *Anti* selective *Sw*HKA (*3S, 4S*) and *syn* selective *Bp*HKA (*3S, 4R*) initially yielded aldol products according to their stereoselectivity at low levels of conversion. Strikingly, both enzymes simultaneously also catalysed the thermodynamic epimerization of the aldol products toward the corresponding equilibrium distribution of *syn* and *anti* configured diastereomers (Figure 2).

Since both *Sw*HKA and *Bp*HKA are highly (*3S*)-selective, *syn*/*anti* epimerization was exclusively observed at the position of the C-4 hydroxyl group. Having made that observation, it is worthy to note that the diastereomeric excess also decreased over time in the case of *Bp*HKA, which is selective for the formation of the thermodynamic product.

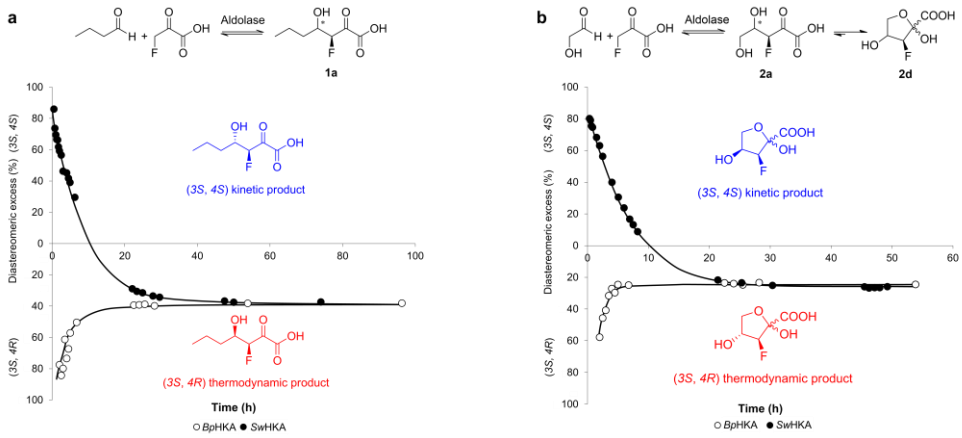


Figure 2: Enzyme catalysed aldol reaction between fluoropyruvate and butyraldehyde **(a)** or glycolaldehyde **(b)**. *Sw*HKA is *anti* selective (*3S*, *4S*), but simultaneously catalyses the epimerization into the *syn* (*3S*, *4R*) configured thermodynamic product. *Bp*HKA is selective for the thermodynamic product (*syn* selective, *3S*, *4R*), but similarly catalyses the epimerization toward the thermodynamic equilibrium distribution of both diastereomers, reducing the diastereomeric excess over time.

***In silico* modelling**

With a straightforward protocol for the study of thermodynamic epimerization effects in hand, we set out to model the impact of different parameters on the catalytic system. Under thermodynamic conditions, the diastereomeric excess is determined by the energy difference between product diastereomers according to equation (1). Strikingly, a $\Delta G \geq 13.2$ kJ/mol is required to achieve a diastereomeric excess of >99% at 298K. Adequate optical purities may therefore not be achievable under thermodynamic conditions for a multitude of molecules.

The stereoselectivity **s** of a catalyst is one of the most important catalytic descriptors, and is derived from the ratio of forward rate constants k_R and k_S, that lead to the corresponding (*R*)- and (*S*)-configured products.[21] For enzymes, the enantioselectivity has been described as *E*,[22] while the diastereoselectivity has not yet been discussed as a separate parameter. In this manner, **s** correlates with the maximal optical purity of the kinetic product. Yet, even highly selective catalysts simultaneously promote product epimerization, and ultimately afford the same diastereomeric excess under thermodynamic conditions (Figure 3a).

The equilibrium constant (K_{eq}) for any type of reaction is set by the temperature and the free energy difference between substrates and products according to equation (1). K_{eq} further correlates with the ratio of apparent rate constants for the forward and reverse reaction (k_f/k_r). The reverse reaction therefore becomes kinetically less significant with an increasing maximal extent of conversion; ultimately converging toward the properties of an irreversible CH/WH type system (Figure 1d, 3b). For bimolecular reactions, the use of one substrate in excess can additionally be applied to drive conversions toward the product side. Based on our model, the ratio of forward to reverse rate constants appears to be the most important parameter for simultaneous product epimerization during the synthesis of diastereomers.

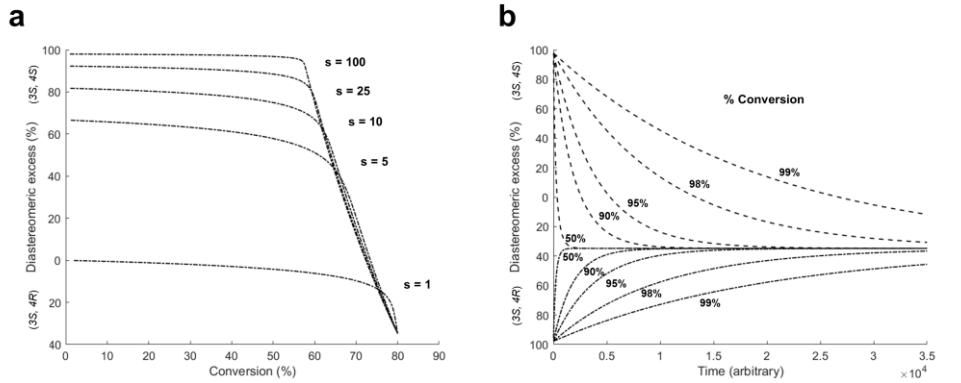


Figure 3: (a) Model for the aldol reaction between fluoropyruvate and butyraldehyde with hypothetical catalysts of different diastereoselectivity **s**, which determines the maximal optical purity of the kinetic product, but even high **s** values do not suppress epimerization. **(b)** Higher levels of conversion (equivalent to a high K_{eq}) increase the ratio of rate constants for the forward and reverse reaction (k_f/k_r), which kinetically disfavours epimerization. The system therefore converges toward CH/WH type features with increasing maximal extent of conversions. Model parameters: ΔG_{RS} = 1.8 kJ/mol, **s** = 100, 298 K.

Experimental validation

In order to validate the results from our model, the catalyst loading was initially varied (Figure 4a). This demonstrated that the observed epimerization of diastereomers was indeed enzyme catalysed and proceeded against the catalyst's stereoselectivity. Notably, the diastereomeric excess of the product mixture was found to be independent of the reaction time and catalyst loading, correlating only with the extent of conversion. This observation has significant implications from both an economic and an environmental perspective, where high conversions are desirable.[23] Maximal conversions involve a concomitant loss of stereocontrol by the catalyst, as the reaction converges toward the thermodynamic equilibrium (Figure 4a). Here, the use of one substrate in excess may remedy this drawback.

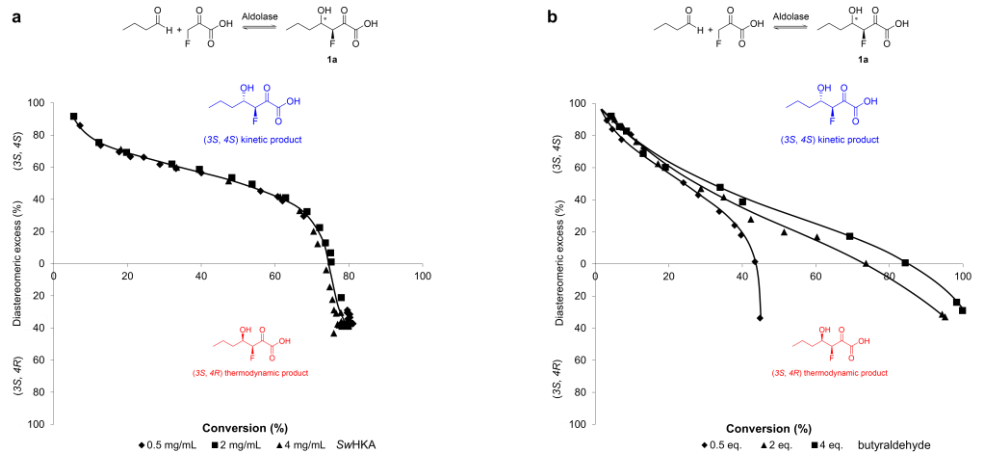


Figure 4: *Sw*HKA catalysed aldol reaction between fluoropyruvate and butyraldehyde. **(a)** The diastereomeric excess correlates with the extent of conversion and is independent of the catalyst loading. **(b)** The use of one substrate in excess (here: butyraldehyde) allows for higher conversions and increases selectivity for the kinetic product.

In agreement with our model, the addition of up to four equivalents of butyraldehyde did not only increase the overall conversion, but also reduced epimerization of the kinetic product (Figure 4b). While solubility limitations prevented us from further increasing the substrate loading, enzymes from other classes, such as transaminases, are commonly applied with one substrate in excess,[24] which can be expected to increase the duration of kinetically controlled conditions accordingly. This was also observed in the aforementioned study of threonine aldolases.[17]

Similar to non-catalytic Diels-Alder reactions, variation of the reaction temperature should allow for a switch from thermodynamically to kinetically controlled conditions and *vice versa*. Indeed, low temperatures were found to increase the selectivity toward formation of the kinetic product, whereas elevated temperatures generally promoted thermodynamic epimerization (Figure 5).

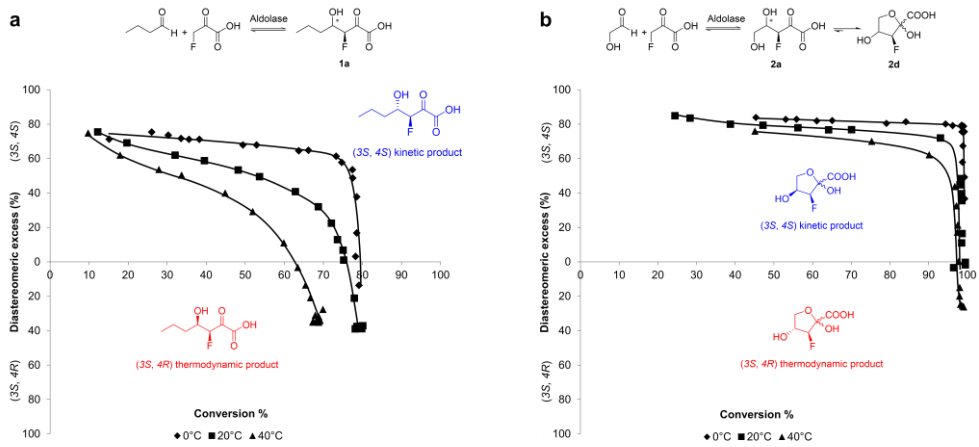


Figure 5. Strategies to promote kinetically controlled conditions. a) Temperature: *Sw*HKA catalysed synthesis of **1a**. At 0°C, the reverse reaction in the retro-aldol direction becomes kinetically negligible, which circumvents product epimerization and allows for higher conversions in the case of exergonic reactions. Conversely, elevated temperatures promote epimerization *via* higher rates in the retro-aldol direction. b) Follow-up reactions (here: ketal formation) increase the overall extent of conversion. This prolongs the duration of kinetically controlled conditions and allows for higher yields of the kinetic product.

Notably, the C-4 stereoselectivity of *Sw*HKA remained similar over the tested temperature range (0 – 40 °C, **s** ≈ 9 corresponding to a *de* ≈ 80%). The increase in selectivity toward formation of the kinetic product (Figure 5a) therefore seems to arise from a significantly lower rate for the reverse reaction in the retro-aldol direction, which is essential for product epimerization. Aldol reactions with glycolaldehyde generally displayed a higher selectivity toward formation of the kinetic product, and the effect of temperature on the extent of epimerization was notable, but less pronounced (Figure 5b). While reactions with aliphatic aldehydes comprise of only one step (namely the aldol reaction, Figure 5a), reactions with hydroxyaldehydes comprise of two consecutive equilibrium reactions (the aldol reaction, followed by a spontaneous intramolecular cyclization, Figure 5b). This

increases the total free energy change of reaction and thereby also the overall extent of conversion. Reaction systems which allow for a high level of conversion by means of follow-up reactions are therefore similarly less prone to thermodynamic epimerization. However, due to the reversible nature of both reactions, complete epimerization still occurs when the overall cascade converges towards the thermodynamic equilibrium for maximal conversions. If the follow-up reaction is rendered irreversible (e.g. *via* selective decarboxylation), a Curtin-Hammett type system is obtained, in which epimerization no longer occurs.[14] Follow-up reactions are therefore a suitable means to increase diastereoselectivity for the kinetic product, and might even be used to render the reaction irreversible in some specific cases.[14]

Quantum mechanical energy calculations

Given the evidence presented, we conclude that thermodynamic contributions can substantially reduce the diastereomeric excess of catalytic reactions and may even lead to the inversion of a stereocentre. To address this issue, thermodynamic analyses are essential to enable rational choices between thermodynamically or kinetically controlled conditions at an early stage of process design. The required relative energies of formation for each diastereomer can be predicted by quantum mechanical (QM) modelling.

To demonstrate this approach, the lowest energy conformations of aldol products **1a** were found through a comprehensive conformational search with DFT. For the resulting conformers, the Gibbs energy of formation was predicted for an aqueous environment using a highly accurate complete basis set method. The computational results correctly predicted the (*3S, 4R*) configuration of the thermodynamic product and were in good agreement with the experimentally observed diastereomeric excess (de_{exp} = 37%, de_{calc} = 55%, Figure 2a).

When considering industrial applications, the two-step synthesis of **2d** constitutes an excellent model system for the study of epimerization effects in sequential equilibrium reactions, such as organocatalytic cascade reactions,[25, 26] cell-

factories[27] and *in vitro* systems biocatalysis[28] (Figure 6). In this system, *Sw*HKA or *Bp*HKA catalyses the aldol reaction between fluoropyruvate and glycolaldehyde to yield the *syn* and *anti* configured diastereomers of **2a**. Spontaneous cyclization subsequently affords both (*2R*) and (*2S*) configured anomers of each diastereomer. Following a similar approach, the free energies of formation were predicted for all diastereomers of **2a** and **2d**. Notably, (*3S*,*4S*) and (*3S,4R*) configured **2a** were predicted to afford a diastereomeric excess of only 2% under thermodynamic conditions. However, subsequent anomerization of **2a** affords (*2R, 3S, 4R*) configured **2d** as the overall thermodynamic product, and the predicted diastereomeric excess of (*3S, 4S*) and (*3S, 4R*) configured anomers was in excellent agreement with experimental results (de_{exp} = 25%, de_{calc} = 21%).

The absolute configuration of the diastereomer that is overall lowest in energy determines the identity of the final product in sequential reactions, which may be opposite to the diastereoselectivity of catalysts that were used for the preceding reaction steps. Mechanistically reversible cascade reactions therefore converge toward the global, thermodynamic product; counteracting both the stereoselectivity of the used catalysts and the relative energies of all reaction intermediates.

The reported error for the quantum chemical prediction of absolute Gibbs energies of formation by the CBS-QB3 method is approximately 4.6 kJ/mol.[29] Since epimers only differ in their absolute configuration at one stereocentre, systematic errors are largely cancelled out in the prediction of *relative* energy differences. In our case, an error of 0.2-0.8 kJ/mol (corresponding to ± 4-16% diastereomeric excess) was observed with respect to experimental values of **1a** and **2d**. It remains to be established whether the low prediction errors are typical for this kind of calculations. The results indicate that CBS-QB3, and QM methods that provide similar accuracy, are promising tools for the prediction of the diastereomeric excess under equilibrium conditions.

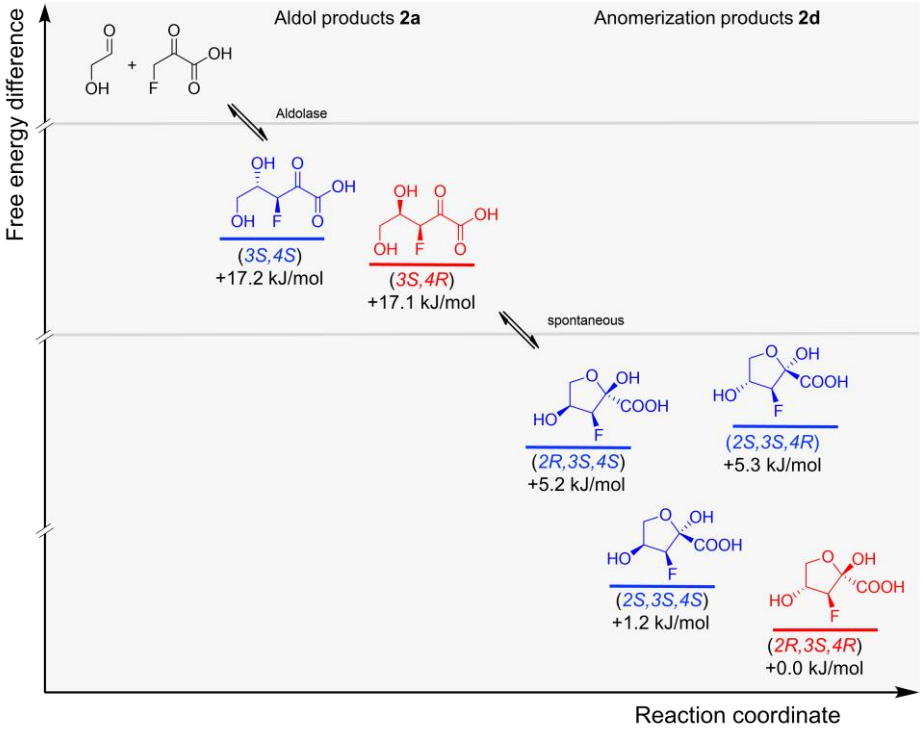


Figure 6. Sequential equilibrium reactions in a two-step cascade system. The aldolase catalysed coupling of fluoropyruvate and glycolaldehyde affords *syn* and *anti* configured **2a**, which spontaneously cyclize to afford the anomerization products **2d**. While **2a** constitutes a nearly racemic product mixture at C-4, extension of the reaction sequence by subsequent cyclization yields the anomers of (*3S, 4R*) configured **2d** as the overall thermodynamic product with a diastereomeric excess of 25%.

8.3 Conclusions

A defining property of catalysts is the increase in rate, by which a reaction converges toward the thermodynamic equilibrium. In agreement with this notion, our results showed that catalysts simultaneously also accelerate the rate of epimerization during the synthesis of diastereomers. This effect can even lead to the inversion of a chiral centre against the catalyst's diastereoselectivity. The common practice of running a reaction to completion and subsequently deducing

the catalyst's diastereoselectivity based on the absolute configuration of the isolated product therefore constitutes a concerning pitfall for catalysis. Predictions of the free energy difference between diastereomers could allow for a rational choice between thermodynamically or kinetically controlled conditions at an early stage of process design. Quantum mechanical calculations with CBS-QB3 were indeed shown to accurately predict both the relative energy difference between the investigated epimers. Kinetically controlled conditions were successfully established by reducing the temperature, but came at the expense of a lower rate of reaction. Notably, a similar effect was achieved by adding one substrate in excess. This did not only increase the extent of conversion, but also improved the selectivity toward formation of the kinetic product. While artificial cascade reactions promise a more sustainable, one-pot synthesis of stereocomplex molecules, the reversible nature of most organocatalytic- or enzymatic reactions inevitably introduces epimerization effects that need to be taken into account.

Acknowledgements

Financial support from the ERA-IB scheme (grant ERA-IB-15-110) to L.M. is gratefully acknowledged. Calculations were performed at the peregrine computer cluster at the University of Groningen.

8.4 Methods Summary

General protocol for aldol reactions to afford 1a-3a. Purified *apo-Sw*HKA or *apo-Bp*HKA was incubated with $MgCl_2$ (2 mM). The aldol reaction was initiated by the addition of sodium fluoropyruvate (50 mM) and the corresponding aldehyde (55 mM) to afford a final volume of 10 mL (5 mM potassium phosphate, 100 mM triethanolamine, pH 7.5). The reaction was monitored by ^{19}F-NMR. Decarboxylation with H_2O_2 was used to analyse both the extent of conversion and the diastereomeric excess, while asymmetric reduction with L-lactic dehydrogenase was used to determine the absolute configurations. Equilibrium conditions were confirmed by the addition of fresh enzyme; no change was observed.

Decarboxylation of the ketoacid aldol products 1a-3a to afford 1b-3b. Samples (400 µL) were quenched by the addition of H_2O_2 (1% *wt/v*) and incubated with D_2O (10% *v/v*, 30 min.,

room temperature) to afford a final volume of 500 µL. Residual H_2O_2 was quenched by the addition of catalase (0.02 mg/mL, 30 min., room temperature). KF was added as an internal standard (2 mM, δ = -120.0 ppm) and the samples were analysed by ^{19}F-NMR. The diastereomeric excess was calculated from the integrals of the peaks corresponding to the (*3S*, *4S*) and (*3S*, *4R*) configured products. Conversions were determined with respect to the consumption of fluoropyruvate.

Reduction of the ketoacid aldol products to afford 1c-3c. Samples (400 µL) were mixed with EDTA (50 mM, 30 min., room temperature). L-lactic dehydrogenase (0.3 mg/mL) was added, using either stoichiometric amounts of NADH (65 mM, 2.0 eq.) or a glucose dehydrogenase (GDH) regeneration system (0.5 mM NADH, 200 mM D-glucose, 0.14 mg/mL GDH) for the asymmetric reduction of the aldol product (overnight, room temperature), after which D_2O (10 % v/v) and KF (2 mM) were added to afford a final volume of 600 µL before analysis by ^{19}F-NMR.

Quantum mechanical modelling. A comprehensive set of conformers was systematically generated for **1a-3a** using canonical dihedral angles in YASARA. The obtained geometries were pre-optimized by the semi-empirical PM3 method using MOPAC Under Gaussian 09, the geometry and energy of each conformer was further optimized using APFD. Frequency calculations were performed to estimate the entropic contributions: Keywords: #p opt=(CalcFC,tight) freq apfd/6-311+g(2d,p) scrf=(iefpcm,solvent=water) Int=UltraFine. All output files were checked for imaginary frequencies, and the lowest energy conformations were selected (defined as those within 15 kJ/mol from the lowest energy structure). These conformations were re-analysed with the more accurate CBS-QB3 method using SMD to model the water environment. Keywords: #p opt freq cbs-qb3 scrf=(solvent=water,smd) geom=checkpoint, at 298.15 K. The predicted Gibbs energies were then used to calculate the energy differences between the respective diastereomers.

References

[1] Smith, S. W., Chiral toxicology: it's the same thing… only different. *Toxicol. Sci.* **2009,** *110* (1), 4-30.

[2] Clayden, J.; Greeves, N.; Warren, S., *Organic Chemistry*, 2nd ed.; OUP Oxford: **2012**.

[3] Marsden, S. R.; Mestrom, L.; McMillan, D. G. G.; Hanefeld, U., Thermodynamically and Kinetically Controlled Reactions in Biocatalysis–from Concepts to Perspectives. *ChemCatChem* **2020,** *12*, 426-437.

[4] Ballard, A.; Narduolo, S.; Ahmad, H. O.; Keymer, N. I.; Asaad, N.; Cosgrove, D. A.; Buurma, N. J.; Leach, A. G., Racemisation in Chemistry and Biology. *Chem. Eur. J.* **2020,** *26*, 3661-3687.

[5] Huerta, F. F.; Minidis, A. B.; Bäckvall, J.-E., Racemisation in asymmetric synthesis. Dynamic kinetic resolution and related processes in enzyme and metal catalysis. *Chem. Soc. Rev.* **2001,** *30* (6), 321-331.

[6] Seeman, J. I., Effect of conformational change on reactivity in organic chemistry. Evaluations, applications, and extensions of Curtin-Hammett Winstein-Holness kinetics. *Chem. Rev.* **1983,** *83* (2), 83-134.

[7] Seeman, J. I., The Curtin-Hammett principle and the Winstein-Holness equation: new definition and recent extensions to classical concepts. *J. Chem. Educ.* **1986,** *63* (1), 42-48.

[8] Nicolaou, K. C.; Snyder, S. A.; Montagnon, T.; Vassilikogiannakis, G., The Diels–Alder reaction in total synthesis. *Angew. Chem. Int. Ed.* **2002,** *41* (10), 1668-1698.

[9] Funel, J. A.; Abele, S., Industrial applications of the Diels–Alder reaction. *Angew. Chem. Int. Ed.* **2013,** *52* (14), 3822-3863.

[10] Vögeli, B.; Erb, T. J., 'Negative'and 'positive catalysis': complementary principles that shape the catalytic landscape of enzymes. *Curr. Opin. Chem. Biol.* **2018,** *47*, 94-100.

[11] Fesko, K.; Reisinger, C.; Steinreiber, J.; Weber, H.; Schürmann, M.; Griengl, H., Four types of threonine aldolases: similarities and differences in kinetics/thermodynamics. *J. Mol. Catal. B: Enzym.* **2008,** *52*, 19-26.

[12] Shibata, K.; Shingu, K.; Vassiley, V. P.; Nishide, K.; Fujita, T.; Node, M.; Kajimoto, T.; Wong, C.-H., Kinetic and thermodynamic control of L-threonine aldolase catalyzed reaction and its application to the synthesis of mycestericin D. *Tetrahedron Lett.* **1996,** *37* (16), 2791-2794.

[13] Kimura, T.; Vassilev, V. P.; Shen, G.-J.; Wong, C.-H., Enzymatic synthesis of β-hydroxy-α-amino acids based on recombinant D-and L-threonine aldolases. *J. Am. Chem. Soc.* **1997,** *119* (49), 11734-11742.

[14] Steinreiber, J.; Schürmann, M.; Wolberg, M.; van Assema, F.; Reisinger, C.; Fesko, K.; Mink, D.; Griengl, H., Overcoming thermodynamic and kinetic limitations of aldolase-catalyzed reactions by applying multienzymatic dynamic kinetic asymmetric transformations. *Angew. Chem. Int. Ed.* **2007,** *46* (10), 1624-1626.

[15] Chen, Q.; Chen, X.; Feng, J.; Wu, Q.; Zhu, D.; Ma, Y., Improving and Inverting Cβ-stereoselectivity of Threonine Aldolase via Substrate-Binding-Guided Mutagenesis and a Stepwise Visual Screening. *ACS Catalysis* **2019,** *9* (5), 4462-4469.

[16] Song, W.; Wang, J.-H.; Wu, J.; Liu, J.; Chen, X.-L.; Liu, L.-M., Asymmetric assembly of high-value α-functionalized organic acids using a biocatalytic chiral-group-resetting process. *Nat. Commun.* **2018,** *9* (1), 3818.

[17] Gutierrez, M. L.; Garrabou, X.; Agosta, E.; Servi, S.; Parella, T.; Joglar, J.; Clapés, P., Serine Hydroxymethyl Transferase from *Streptococcus thermophilus* and L-threonine aldolase from *Escherichia coli* as Stereocomplementary Biocatalysts for the Synthesis of β-hydroxy-α,ω-diamino Acid Derivatives. *Chem. Eur. J.* **2008**, *14*, 4647-4656.

[18] Hernandez, K.; Zelen, I.; Petrillo, G.; Usón, I.; Wandtke, C. M.; Bujons, J.; Joglar, J.; Parella, T.; Clapés, P. Engineered L-Serine Hydroxymethyltransferase from *Streptococcus thermophilus* for the Synthesis of α,α-Dialkyl-α-Amino Acids. *Angew. Chem. Int. Ed.* **2015**, *54*, 3013-3017.

[19] Fang, J.; Hait, D.; Head-Gordon, M.; Chang, M. C., Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases. *Angew. Chem. Int. Ed.* **2019,** *58*, 11841-11845.

[20] De Berardinis, V.; Guérard-Hélaine, C.; Darii, E.; Bastard, K.; Hélaine, V.; Mariage, A.; Petit, J.-L.; Poupard, N.; Sánchez-Moreno, I.; Stam, M.; Gefflaut, T.; Salanoubat, M.; Lemaire, M., Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chem.* **2017,** *19* (2), 519-526.

[21] Gawley, R. E., Do the Terms "% ee" and "% de" Make Sense as Expressions of Stereoisomer Composition or Stereoselectivity? *J. Org. Chem.* **2006**, *71*, 2411-2416.

[22] Chen, C. S.; Wu, S. H.; Girdaukas, G.; Sih, C. J., Quantitative analyses of biochemical kinetic resolution of enantiomers. 2. Enzyme-catalyzed esterifications in water-organic solvent biphasic systems. *J. Am. Chem. Soc.* **1987**, *109*, (9), 2812-2817.

[23] Sheldon, R. A., The E factor 25 years on: the rise of green chemistry and sustainability. *Green Chem.* **2017,** *19* (1), 18-43.

[24] Dunbabin, A.; Subrizi, F.; Ward, J. M.; Sheppard, T. D.; Hailes, H. C., Furfurylamines from biomass: transaminase catalysed upgrading of furfurals. *Green Chem.* **2017,** *19* (2), 397-404.

[25] Grondal, C.; Jeanty, M.; Enders, D., Organocatalytic cascade reactions as a new tool in total synthesis. *Nat. Chem.* **2010,** *2* (3), 167-178.

[26] van Der Helm, M. P.; Klemm, B.; Eelkema, R., Organocatalysis in aqueous media. *Nat. Rev. Chem.* **2019,** *3* (8), 491-508.

[27] Lee, S. Y.; Mattanovich, D.; Villaverde, A., Systems metabolic engineering, industrial biotechnology and microbial cell factories. *Microb. Cell Fac.* **2012,** *11*, 156.

[28] Fessner, W.-D., Systems Biocatalysis: Development and engineering of cell-free "artificial metabolisms" for preparative multi-enzymatic synthesis. *New Biotechnol.* **2015,** *32* (6), 658-664.

[29] Montgomery Jr, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A., A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **2000,** *112* (15), 6532-6542.

9

Conclusions and Outlook

This thesis aimed at contributing to the advancement of our fundamental understanding of different enzymatic modes of operation. A particular focus was placed on the role of thermodynamics, next to the study of specific active site interactions. The transient responses of catalytic systems towards external stimuli and their ultimate convergence towards the thermodynamic equilibrium forms the red line that is common to all chapters in this thesis.

In chapter 2, the mechanistic reversibility of reactions that are catalysed by ThDP-dependent enzymes was demonstrated. The corresponding decarboxylation driven reactions are therefore not irreversible, but kinetically controlled. While this distinction may sometimes seem insignificant from a practical (synthetic) viewpoint, a correct description of the catalytic system at hand is an essential prerequisite for further innovation. For instance, an amperometric biosensor for ThDP was developed based on the reversible splitting of L-erythrulose by transketolase.[1]

Intuition would suggest that a matching polarity is an important requirement for the accommodation of substrates within an enzyme's active site. Yet, the design of interactions between substrates and protein side chains of opposite polarity afforded the most efficient enzyme variants for the conversion of non-natural substrates in chapter 2 (hydrogen bonds between aliphatic aldehydes and a glutamate residue) and chapter 4 (CH-π interactions between hydroxypyruvate and phenylalanine/tyrosine residues). CH-π interactions in particular are currently an underappreciated design feature for substrate

recognition and activation in enzymes beyond the well-known classes of carbohydrate-active enzymes.

Regulation of enzyme activity is a key challenge with respect to selectivity in many multi-step cascade reactions, such as artificial metabolic pathways. In chapter 5, we demonstrated the reversibly activation of a metalloenzyme by the characteristic structural motif of its substrate. Different energies of formation allow for the dynamic movement of the metal cofactor between distinct metal binding sites by ligand exchange. Due to its conceptual simplicity, we anticipate the prevalence of this regulatory mechanism in other enzyme classes, where it might have been overlooked in protein crystal structures until now.

In chapter 7, features of kinetic reaction control as a means for improved product yields were conceptually transferred from the well-studied examples of amidohydrolases and acyl transferases on to other enzyme classes of analogous mechanisms. This review highlights, how in-depth knowledge of an enzyme's mode of operation can lead in a next step to a more encompassing understanding of enzyme catalysis.

In chapter 8, experimental and theoretical evidence was presented for the simultaneous occurrence of thermodynamic epimerization during the catalytic synthesis of diastereomers. These findings warrant the introduction of a new standard for the characterisation of catalysts. Most importantly, thermodynamic epimerization was identified as an important factor to consider for multi-step cascade reactions, which promise substantial reductions in the environmental impact of chemical syntheses.

To conclude, this thesis highlighted the simultaneous competition between thermodynamically- and kinetically controlled features in enzyme catalysis.

Kinetic reaction control is an efficient means to either afford optically pure products in chiral synthesis, or to improve product yields *via* the elimination of a good leaving group. While these concepts are typically treated separately, they share the same underlying principle of competition between a partial (i.e. the kinetic) and a global (i.e. the thermodynamic) equilibrium system. Therefore, they should follow the same mathematical equations and allow for mutual knowledge transfer in the future. For instance, the synthesis to hydrolysis ratio is a particularly well-established parameter in amidohydrolases and acyl transferases, whereas an analogous parameter for asymmetric synthesis is still lacking.

References

[1] M. Halma, B. Doumèche, L. Hecquet, V. Prévot, C. Mousty, F.Charmantray, *Biosensors and Bioelectronics,* **2017**, *87*, 850-857.

Acknowledgements

Looking back at my PhD journey, it is not the completion of this book, but the many ways in which colleagues, family and friends have influenced my work, that gives me the greatest feeling of accomplishment.

First and foremost, I would like to thank my promotor **Ulf Hanefeld**. Initially, I had enrolled for the Biochemical Engineering track of LST, but after hearing your first lecture I already knew that my true passion lay in Biocatalysis. Over the past five years I got to know you as a great scientist, but more importantly also as a kind and dedicated person, who always put his student's interests above everything else. Your door was always open, and it was your trust and support that allowed me to become an independent scientist. I will always be grateful for the efforts that you took to offer me this position, and I am proud to say that I could not have asked for a better supervisor.

I would also like to sincerely thank my co-promotor **Duncan McMillan**. During our first (heavily jetlagged) conversation directly after your flight from New Zealand I knew that our relationship was off to a good start. To me, you will always be the man with a plan for a plan. Life is a game of chess, and you taught me to think not only one, but ten steps ahead. Your advice helped me with navigating the many struggles that inevitably come with doing a PhD.

I would like to thank **Isabel Bento** for joining this project and providing us with much needed structural information. Your expertise allowed us to explore entirely new directions for this project, and I am grateful for your kind hospitality during my stay at the EMBL.

I would also like to thank **Hein Wijma** for his support with quantum mechanical calculations. Your help revealed new insights into the underlying mechanisms, and your input has been of great value during this project.

I would like to thank **Peter-Leon Hagedoorn** for his expertise and support in the analysis of metalloenzymes and for valuable discussions on enzyme kinetics.

I would like to particularly thank **Helen Hailes** for her kind hospitality and a warm welcome to her group during my research stay at UCL. I would also like to thank **John Ward**, **Leona Leipold**, **Fabiana Subrizi** and **Max Cardenás-Fernández** for their help and advice in the lab, and all other group members for an amazing time!

Special thanks go out to my paranymphs **Luuk** and **Hugo**. **Luuk**, without you, my PhD experience would not have been the same. You were my toughest reviewer, and our many scientific discussions and brainstorm sessions were invaluable contributions to this thesis. **Hugo**, you are a true (smaken-)wizard, and your magic spiced up not only our craft beers, but also this thesis. It was my friendship with you two that has made Delft my new home, and I've tremendously enjoyed all the Sjpasj that we had (and still have) together!

Florence, you are an amazing housemate and our end-of-the-day happy hours in balconia keep bringing the perfect work-life balance to the home office.

To the PhD committee **Albert, Chema, Marta, Anna, Britte** and **Florence**: I had an amazing time with you during the organisation of the BT symposium, Halloween Posters, dancing workshops… you guys are the best!

I would like to thank my students **Fenneke, Tobias, Meera, Michael** and **Sander** for all of your contributions to my work. Being your supervisor was a unique experience and I am proud of what we have accomplished together.

Unforgotten are also the many great trips to conferences together with **Luuk, Hanna, José, Sébastien, Marine, Georg, Morten, Fabio, Eman, Linda, Natália** and **Alexandra**. It was a pleasure to be your colleague!

Wijb, thanks for all the good moments with brewing beer and game nights!

Sam, while our time as colleagues was unfortunately cut short, I am glad that I was there to witness the demonstration of the rule of thumb in the lab.

Many thanks also go out to **Marc, Remco, Lloyd, Laura** and **Stephen**. Your experience with operating the equipment and your technical support have spared me one or the other grey hair during my PhD!

I would also like to thank **Mieke** for all her help with keeping a project running and mastering the paperwork that comes with doing a PhD.

At last, I would also like to thank all of my (former) colleagues **Frank, Kristina, Caroline, Carolin, Retna, Lorina, Paula, Rosario, Jonathan, Tiago, Mieke, Sabry, Wuyuan, Sandy, Florian, Jia Jia, Elena, Milja** and **Maarten,** who made my time in BOC so enjoyable.

Finally, I would like to express my sincere gratitude to my whole family. Going abroad is never easy, and I am grateful for all of your support with my decision. Thank you for always being there for me and providing a home that I can come back to. Without your guidance, I would not have been the person that I am today.

Curriculum Vitae

Stefan Robert Marsden was born on the 8th of September 1991 in Baden (CH). Growing up in Switzerland, his British and German citizenship laid the foundation for the development of a European identity. He received his Abitur from the Hochrhein-Gymnasium in Waldshut-Tiengen (D), where he graduated within the top 3% of his year and obtained the Gesellschaft Deutscher Chemiker award. He received his BSc in Chemistry from the University of Basel, where he worked on the site-selective modification of DNA / RNA in the group of Prof. D. Gillingham. He continued with the MSc programme Life Science & Technology at Delft University of Technology (NL), where he was involved in the Board of Studies. As a member of a TU Delft *dreamteam*, he won the prestigious iGEM competition in 2015 and received the Royal IHC Teamwork Award. His thesis project in the group of Prof. U. Hanefeld revealed a new understanding of transketolase catalysed reactions, and he graduated from the MSc programme with *Honours* and *Cum Laude.* He received the shared Student of the Year Award from the Faculty of Applied Sciences at TU Delft and was nominated for the KNCV Golden Master Award. He then carried out an industrial internship at BioscienZ (NL), where he co-invented a patent on fungal protein expression. Afterwards, he pursued his PhD under the supervision of Prof. U. Hanefeld and Dr. D. G. G. McMillan, investigating thermodynamic effects in biocatalysis. During this time, he was involved in the PhD committee and organised a departmental symposium. Over the course of his PhD project, he initiated collaborations with Dr. I. Bento from the EMBL Hamburg (D) and Prof. H. C. Hailes from the University College London (UK), where he both worked as a guest researcher, next to a collaboration with Dr. H. J. Wijma from the Rijksuniversiteit Groningen (NL).

Publications and Patents

- ‘Substrate Induced Movement of the Metal Cofactor between Active and Resting State Configurations in a Metalloenzyme’
  Stefan R. Marsden, Hein J. Wijma, Michael Mohr, Peter-Leon Hagedoorn, Jesper Laustsen, Cy M. Jeffries, Luuk Mestrom, Duncan G. G. McMillan, Isabel Bento, Ulf Hanefeld, *manuscript in preparation*.

- ‘Thermodynamics Determine the Diastereochemical Outcome of Catalytic Reactions’
  Stefan R. Marsden, Luuk Mestrom, Hein J. Wijma, Sander J. Noordam, Duncan G. G. McMillan, Ulf Hanefeld, *ChemCatChem,* **2021**. DOI: 10.1002/cctc.202100178.

- ‘Assessing the Thiamine Diphosphate Dependent Pyruvate Dehydrogenase E1 Subunit for Carboligation Reactions with Aliphatic Ketoacids’
  Stefan R. Marsden, Duncan G. G. McMillan, Ulf Hanefeld, *Int. J. Mol. Sci.,* **2020**, *21*, (22), 8641. DOI: 10.3390/ijms21228641.

- ‘Discovery and Engineering of an Aldehyde Tolerant 2-deoxy-D-ribose 5-phosphate Aldolase (DERA) from *Pectobacterium atrosepticum*’
  Meera Haridas, Carolin Bisterfeld, Le Min Chen, **Stefan R. Marsden**, Fabio Tonin, Rosario Médici, Adolfo Iribarren, Elizabeth Lewkowicz, Peter-Leon Hagedoorn, Ulf Hanefeld, Eman Abdelraheem, *Catalysts,* **2020**, *10*, 883. DOI: 10.3390/catal10080883.

- ‘Anomeric selectivity of Trehalose Transferase with Rare L-sugars’
  Luuk Mestrom, **Stefan R. Marsden**, Hessel van der Eijk, Jesper U. Laustsen, Cy M. Jeffries, Dmitri I. Svergun, Peter-Leon Hagedoorn, Isabel Bento, Ulf Hanefeld, *ACS Catalysis*, **2020**, *10*, 8835-8839.
  DOI: 10.1021/acscatal.0c02117

- ‘Comparison of Enzymes immobilized on Immobeads and Inclusion Bodies: A Case Study of Trehalose Transferase’
  Luuk Mestrom, **Stefan R. Marsden**, Duncan G. G. McMillan, Rob Schoevaart, Peter-Leon Hagedoorn, Ulf Hanefeld, *ChemCatChem*, **2020**, *12*, 3249-3256. DOI: 10.1002/cctc.202000241.

- ‘Expression of Ovalbumin and its Natural Variants’
  Joyce J. E. Arnouts, María Vázquez Vitali, Niels Lauret, **Stefan R. Marsden**, Lucie Parenicová, Wilhelmus T. A. M de Laat, *Patent No.* 20151815.6 – 1118.

- ‘Thermodynamically and Kinetically Controlled Reactions in Biocatalysis - From Concepts to Perspectives’
  Stefan R. Marsden, Luuk Mestrom, Duncan G. G. McMillan, Ulf Hanefeld, *ChemCatChem*, **2020**, *12*, 426-437. DOI: 10.1002/cctc201901589.

- ‘LeLoir Glycosyltransferases in Applied Biocatalysis: A miltidisciplinary Approach’
  Luuk Mestrom, Marta Przypis, Daria Kowalczykiewicz, André Pollender, Antje Kumpf, **Stefan R. Marsden**, Isabel Bento, Andrzej B. Jarzębski, Katarzyna Szymańska, Arkadiusz Chruściel, Dirk Tischler, Rob Schoevaart, Ulf Hanefeld, Peter-Leon Hagedoorn, *Int. J. Mol. Sci.,* **2019**, *20*, 5263. DOI: 10.3390/ijms20215263.

- ‘CH-π Interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase’
  Stefan R. Marsden, Luuk Mestrom, Isabel Bento, Peter-Leon Hagedoorn, Duncan G. G. McMillan, Ulf Hanefeld, *Adv. Synth. Catal.*, **2019**, *361*, 2649-2658. DOI: 10.1002/adsc.201900205.

- ‘Artificial Fusion of mCherry Enhanced Solubility and Stability of Trehalose Transferase’
  Luuk Mestrom, **Stefan R. Marsden,** Marit Dieters, Puck Achterberg, Lysanne Stolk, Isabel Bento, Ulf Hanefeld, Peter-Leon Hagedoorn, *Appl. Environ. Microbiol.*, **2019**, *85*, e03084-03018. DOI: 10.1128/AEM.03084-18.

- ‘Correction for Artificial Fusion of mCherry Enhanced Solubility and Stability of Trehalose Transferase’
  Luuk Mestrom, **Stefan R. Marsden,** Marit Dieters, Puck Achterberg, Lysanne Stolk, Isabel Bento, Ulf Hanefeld, Peter-Leon Hagedoorn, *Appl. Environ. Microbiol.*, **2019**, *85*, e00942. DOI: 10.1128/AEM.03084-18.

- ‘Separating Thermodynamics from Kinetics – A New Understanding of the Transketolase Reaction’
  Stefan R. Marsden, Lorina Gjonaj, Stephen J. Eustace, Ulf Hanefeld, *ChemCatChem,* **2017**, *9*, 1808-1814. DOI: 10.1002/cctc.201601649.