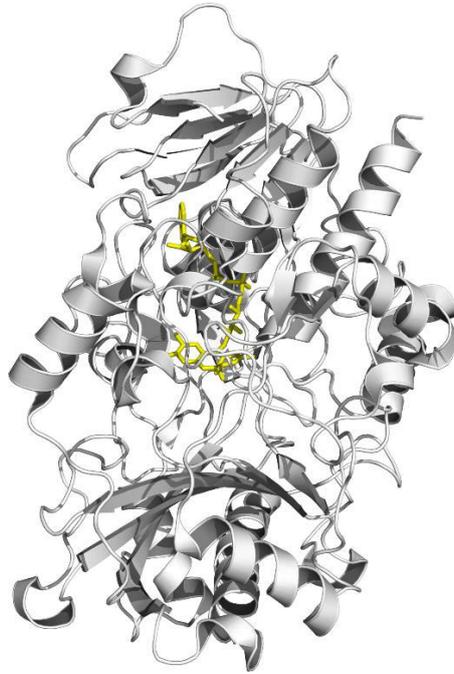


Probing the active site of pyranose dehydrogenase by experimental and computational means



PhD Thesis

DI Michael Graf

Vienna, December 2014

Supervisor: Univ. Prof. Dietmar Haltrich

Food Biotechnology Laboratory

Department of Food Science and Technology

BOKU – University of Natural Resources and Life Sciences

Muthgasse 11, 1190 Vienna, Austria



*"Das schönste Glück des denkenden Menschen ist,
das Erforschliche erforscht zu haben und das
Unerforschliche ruhig zu verehren."*

Johann Wolfgang von Goethe

Table of Contents

Abstract	vii
Kurzfassung	ix
Chapter 1	1
<hr/>	
Introduction	
Chapter 2	25
<hr/>	
Aim and Outline	
Chapter 3	29
<hr/>	
D-Glucose Dioxidation by PDH	
Chapter 4	51
<hr/>	
Ligand Promiscuity of PDH	
Chapter 5	87
<hr/>	
PDH – Experiments and Simulations	
Chapter 6	139
<hr/>	
Conclusion and Outlook	
List of Publications	144
CV	145
Acknowledgements	146

Abstract

Pyranose dehydrogenase from *Agaricus meleagris* (*Am*PDH), a member of the structural family of GMC oxidoreductases, carries a monocovalently linked FAD cofactor. *Am*PDH (di)oxidizes many different sugars at C1–C4 and is potentially involved in lignin degradation. It is inactive with dioxygen but uses quinones and organometallic compounds instead in its oxidative half-reaction. Functional assignments of active site residues in the enzyme are still elusive to date, making rational protein engineering approaches towards applications in sugar conversions, organic synthesis, and bioelectrochemistry difficult. In the course of this thesis, the monocovalent FAD-linkage and the vicinity of the isoalloxazine in the active site of *Am*PDH were investigated. Wild type *Am*PDH and active site variants were heterologously expressed in the yeast *Pichia pastoris*, purified, and characterized by biochemical, biophysical, and computational means. Experimental techniques included amongst others ECD, TCA/acetone precipitation, GC-MS of reaction products, as well as measuring the oxygen reactivity and redox potential. UV-Vis and stopped-flow spectroscopy were utilized to elucidate steady-state and pre-steady-state kinetics, respectively. Molecular dynamics (MD) simulations and free energy calculations rationalized or suggested experiments.

Chapter 1 of this thesis provides an overview of flavin-dependent enzymes, the structural family of GMC oxidoreductase and selected members thereof. Moreover, a short introduction about (computational) enzyme engineering and case studies of selected GMC oxidoreductases are given.

Chapter “3” presents the first study employing MD simulations on *Am*PDH. The thermodynamic driving forces for D-glucose binding and important interactions between the enzyme and its substrate were investigated. The outcomes of this chapter suggested targets for site directed mutagenesis, which are dealt with in detail in **Chapter 5**. **Chapter 4** describes extensive MD simulations on *Am*PDH with a bound sugar substrate, which rationalizes the substrate promiscuity of the enzyme by combining free energy calculations and reweighted distance analysis. In **Chapter 5**, both biochemical and biophysical investigations were performed to elucidate structure–function relationships of active site residues in *Am*PDH. MD simulations corroborated the experimental findings and gave additional insights in dynamic properties. **Chapter 6** proposes a conclusion and outlook based on the results obtained during this PhD thesis.

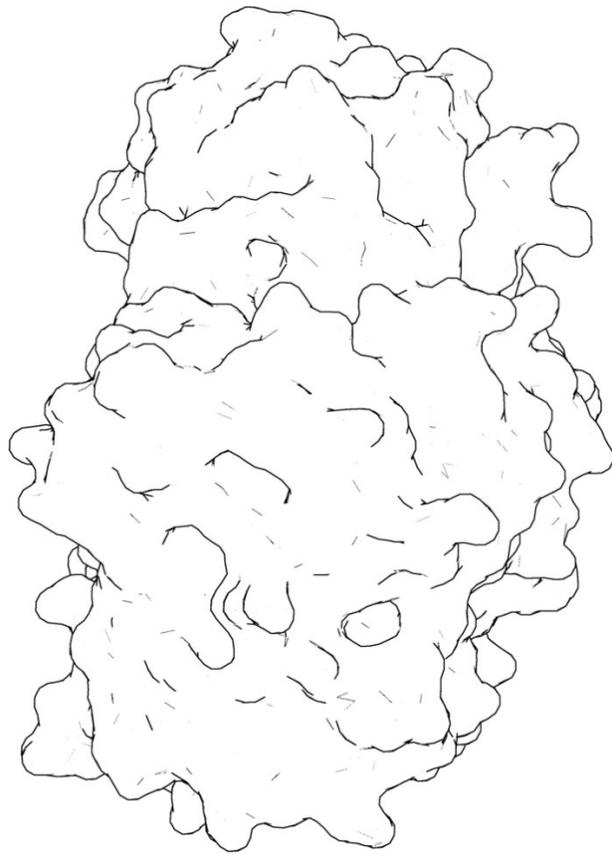
Kurzfassung

Die GMC Oxidoreduktase Pyranose Dehydrogenase von *Agaricus meleagris* (*Am*PDH) beinhaltet einen monokovalent gebundenen FAD Cofaktor. *Am*PDH (di)oxidiert viele verschiedene Zucker an Position C1–C4 und ist vermutlich im Ligninabbau involviert. Das Enzym reagiert nicht mit Sauerstoff, dafür aber mit Chinonen und organometallischen Verbindungen. Die Struktur–Funktions Beziehungen der Aminosäuren im aktiven Zentrum von *Am*PDH wurden bisher noch nicht eingehend untersucht. Dadurch sind gezielte Veränderungen, um das Enzym für Zucker-Derivatisierungen, organische Synthesen und in der Bioelektrochemie zu verwenden, schwierig. Im Zuge dieser Dissertation wurden die kovalente Bindung zum FAD und die unmittelbare Umgebung des Isoalloxazins im aktiven Zentrum von *Am*PDH untersucht. Wildtyp *Am*PDH und Mutanten des aktiven Zentrums wurden heterolog in der Hefe *Pichia pastoris* exprimiert und anschließend gereinigt. Danach wurden die Enzyme biochemisch, biophysikalisch und computergestützt charakterisiert. Die experimentellen Methoden umfassten unter anderem ECD, TCA/Aceton Präzipitation, GC-MS der Reaktionsprodukte sowie die Bestimmung der Sauerstoffreaktivität und des Redoxpotentials. Die Steady-State und Pre-Steady-State Kinetik wurden mittels UV-Vis beziehungsweise Stopped-Flow-Spektroskopie bestimmt. Molekulardynamik (MD) Simulationen und Berechnungen der freien Energie konnten experimentelle Ergebnisse untermauern oder anregen.

Kapitel 2 dieser Dissertation gibt einen Überblick über Flavoenzyme und GMC Oxidoreduktasen inklusive ausgewählter Vertreter davon. Weiters wird (computergestütztes) Enzym-Engineering allgemein beschrieben und mittels Fallbeispielen von GMC Oxidoreduktasen näher erläutert. **Kapitel 3** beschreibt die bisher erste Untersuchung von *Am*PDH, die mittels MD Simulationen durchgeführt wurde. Die thermodynamischen Grundlagen der Glucosebindung sowie wichtige Interaktionen zwischen Glucose und *Am*PDH wurden bestimmt. Die Ergebnisse dieses Kapitels führten zu vielversprechenden Vorschlägen, die in nachfolgenden Experimenten mittels ortsspezifischer Mutagenese in **Kapitel 5** durchgeführt wurden. **Kapitel 4** beschreibt umfangreiche MD Simulationen an *Am*PDH mit gebundenem Zuckersubstrat. Durch Berechnungen der freien Energie in Kombination mit gewichteter Distanzanalyse konnte die Substrat-Promiskuität des Enzyms erklärt werden. In **Kapitel 5** wurden umfangreiche experimentelle Studien zur Struktur–Funktions Beziehung von Aminosäuren im aktiven Zentrum von *Am*PDH durchgeführt. Auch hier konnten MD Simulationen experimentelle Ergebnisse untermauern und zusätzliche Erkenntnisse liefern. In **Kapitel 6** werden einige Schlussfolgerungen gezogen und ein Ausblick in die Zukunft gegeben.

CHAPTER 1

Introduction



Author

Michael Graf

1.1 Flavin-dependent enzymes

Flavin-dependent enzymes contain tightly bound FMN or FAD as cofactor, which are both composed of riboflavin (vitamin B₂) as well as a phosphate group (in FMN) or adenine diphosphate (ADP) in FAD (**Figure 1A**) [1]. In either case, the cofactor comprises an isoalloxazine ring, which is the redox-active center of these enzymes. The cofactor is yellow in the oxidized state and colorless in the reduced state (**Figure 1B**), making flavoproteins readily accessible to spectroscopic studies [2]. Concentrations of FAD and FMN free in solution can be determined *via* the known extinction coefficients for the oxidized species at their absorption maxima [3]: $\epsilon_{450} = 11,300 \text{ M}^{-1} \text{ cm}^{-1}$ for FAD [4] and $\epsilon_{446} = 12,200 \text{ M}^{-1} \text{ cm}^{-1}$ for FMN [5]. This allows to determine the extinction coefficient for a flavin bound to a protein, which usually causes a hypsochromic or bathochromic shift of the absorption maximum around 450 nm.

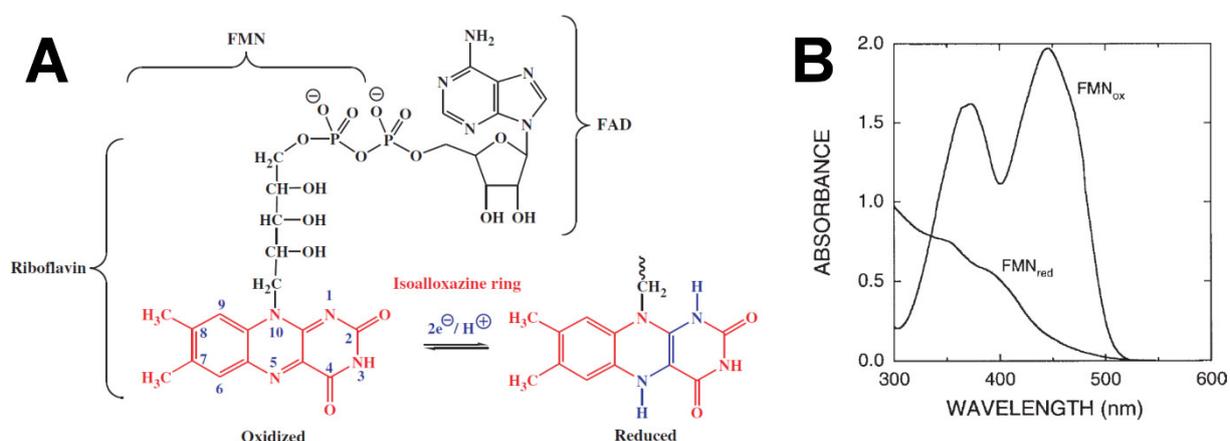


Figure 1: (A) Structures of riboflavin, FMN, and FAD. The redox-active isoalloxazine ring is depicted in red (oxidized) and blue (two-electron reduced state) and the numbering scheme for the isoalloxazine ring in blue. From Macheroux and coworkers [6]. (B) UV-visible spectrum of FMN in its oxidized (FMN_{ox}) and reduced (FMN_{red}) state. From Macheroux [2].

As already indicated in **Figure 1A**, the flavin moiety – or more precisely the isoalloxazine ring – can be reversibly reduced and oxidized. Thereby, two half reactions can be distinguished in flavoenzyme catalysis: the reductive and the oxidative half reaction (**Figure 2A**) [7]. Despite its central role in these redox reactions, the flavin is not irreversibly altered but acts as an electron shuttle. Fully reduced or oxidized flavins are accessible *via* the one-electron reduced semiquinone (**Figure 2B**). These semiquinones have distinct peaks in UV-visible spectra, and are therefore readily detectable as well [8,9]. The spectral properties, corresponding to certain redox intermediates, make flavoproteins one of the best-studied enzyme families to date [10,11].

The above mentioned redox processes occur mainly at two sites of the isoalloxazine ring: the N5-atom and the C4a-atom (atom numbers in **Figure 1A** or **Figure 2A**): additions of hydrides [12,13]

and carbanions [14] occur *via* N5, while covalent links with thiols [15] and radical oxygen species [16] occur at C4a [7] (**Figure 2C**).

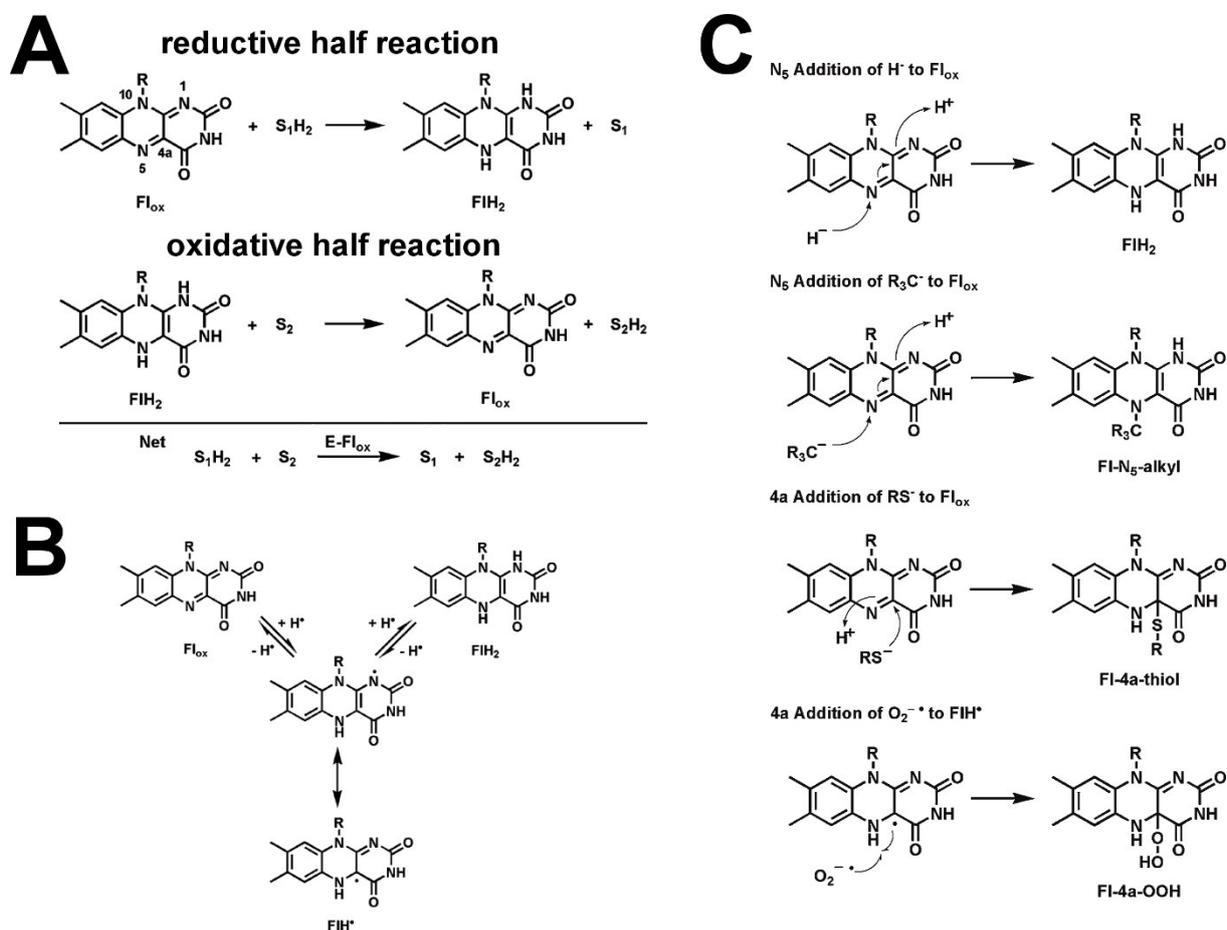


Figure 2: (A) The two redox half reactions in flavoenzyme catalysis: reductive- and oxidative half reaction. (B) Kinetically and thermodynamically accessible flavin redox states: oxidized flavin (Fl_{ox}), one-electron reduced semiquinone (FlH^{\bullet}), and two-electron reduced dihydroflavin (FlH_2). (C) N5 and C4a as electron entry and exit sites from the isoalloxazine ring: covalent 4a-thiol and 4a-hydroperoxide adducts. From Walsh and Wencwicz [7].

About 10% of all flavoenzymes carry a mono- or bicovalently tethered cofactor, which is linked *via* a cysteine, histidine, threonine, or tyrosine to the protein, whereas the flavin is attached *via* its C8-methyl or C6-atom of the isoalloxazine or a phosphate group involved in a phosphoester bond (**Figure 3**) [17]. Covalent flavinylation is an autocatalytic process [18,19], upon which the redox potential and the protein stability are usually increased [17]. Additionally, covalent flavin attachment is thought to protect the flavin from unwanted modifications or inactivation and to tune its reactivity for electron transfer or oxygen reactivity [1,17].

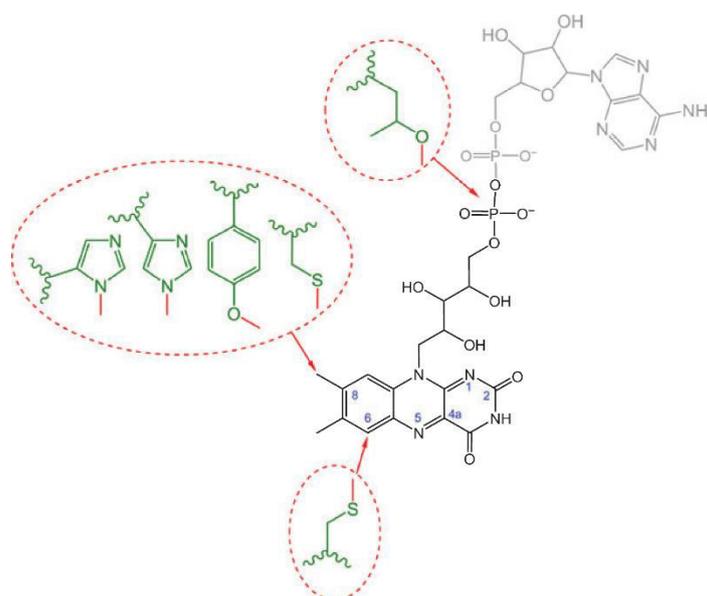


Figure 3: Types of covalent flavin–protein linkages known to date. Color code: FMN (black), FAD (black and gray), and linking amino acids (green). Red arrows indicate covalent attachment sites and blue numbers selected isoalloxazine atoms. From Heuts and coworkers [17].

The features described above make flavoenzymes extremely versatile catalysts, which take part in many different biological processes (**Figure 4A**) [1,7,11]. Consequently, flavoproteins from five out of the six enzyme classes are known, with hydrolases as the only class of which no representative is known to date (**Figure 4B**) [6]. Since the flavin cofactor is able to adopt so many redox intermediates, it is not surprising that 91% of all known flavoenzymes are oxidoreductases, though.

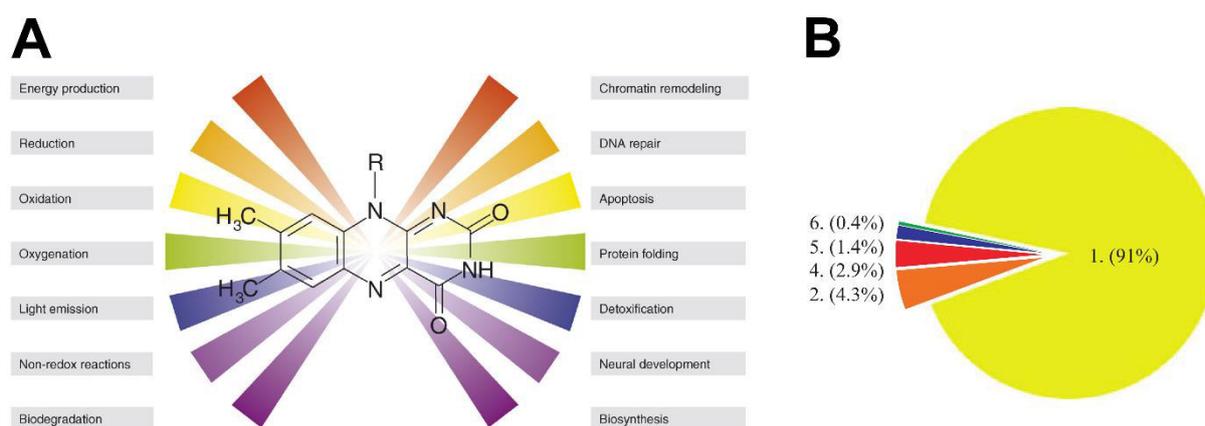


Figure 4: (A) Biological functions of flavoenzymes. From Joosten and van Berkel [1]. (B) Flavoproteins belonging to various enzyme classes. Color code: class 1, oxidoreductases (yellow); class 2, transferases (orange); class 4, lyases (red); class 5, isomerases (blue); class 6, ligases (green). From Macheroux and coworkers [6].

According to the numbering scheme of the ‘International Union of Biochemistry and Molecular Biology’ (IUBMB), oxidoreductases form the first enzyme class (**Figure 4B**). Oxidoreductases are further grouped into subclasses, which indicate the chemical species that is attacked (*e.g.* alcohols, aldehydes, or amines), to eventually transfer electrons from a reductant to an oxidant. This feature makes this enzyme class attractive for biotechnological applications, because many industrial processes involve redox reaction, such as the production of specialty chemicals, pharmaceuticals, and food additives. More general features of enzymes, which usually catalyze reactions with great selectivity and stereospecificity and efficiently work under relatively mild conditions are further incentives to utilize oxidoreductases in the described contexts [20]. Another field of application is bioelectrochemistry, where an electric current is generated by these enzymes, which is then used to power biofuel cells or generate a signal for a biosensor [21–23]. The previously mentioned versatility of the flavin cofactor makes flavin-dependent oxidoreductases especially promising catalysts for these applications.

1.2 GMC Oxidoreductases

The structural family of glucose–methanol–choline (GMC) oxidoreductases has first been described by Cavener in 1991 and its members can be found in pro- and eukaryotes [24]. The common structural features of this family are the PHBH (*p*-hydroxybenzoate hydroxylase) like fold [25], the flavin binding subdomain comprising an ADP-binding $\beta\alpha\beta$ motif (Rossmann fold) close to their N-terminus [24], and a substrate binding subdomain [26,27]. The substrate binding subdomain is less conserved throughout the GMC oxidoreductases, which probably reflects the different substrate specificities of the family members, ranging from mono- and polysaccharides to cholesterol [28]. Phylogenetic analysis of the GMC-superfamily of flavoproteins clearly revealed its main clades (**Figure 5**) [28,29].

The focus of this PhD thesis was an in-depth investigation of the active site in pyranose dehydrogenase from *Agaricus meleagris* (PDH, EC 1.1.99.29) by experimental and computational means. Where possible, results were compared to the GMC family members aryl-alcohol oxidase (AAO, EC 1.1.3.7), choline oxidase (CHO, EC 1.1.3.17), glucose 1-oxidase (GOX, EC 1.1.3.4), and pyranose 2-oxidase (P2O, EC 1.1.3.10). In the next sections, key features of PDH as well as AAO, CHO, GOX, and P2O are introduced. Owing to the focus of the thesis, PDH is described in most detail. For further information, the reader is directed to the references given in the text as well as a recent review about the substrate oxidation mechanism of several GMC family members [30].

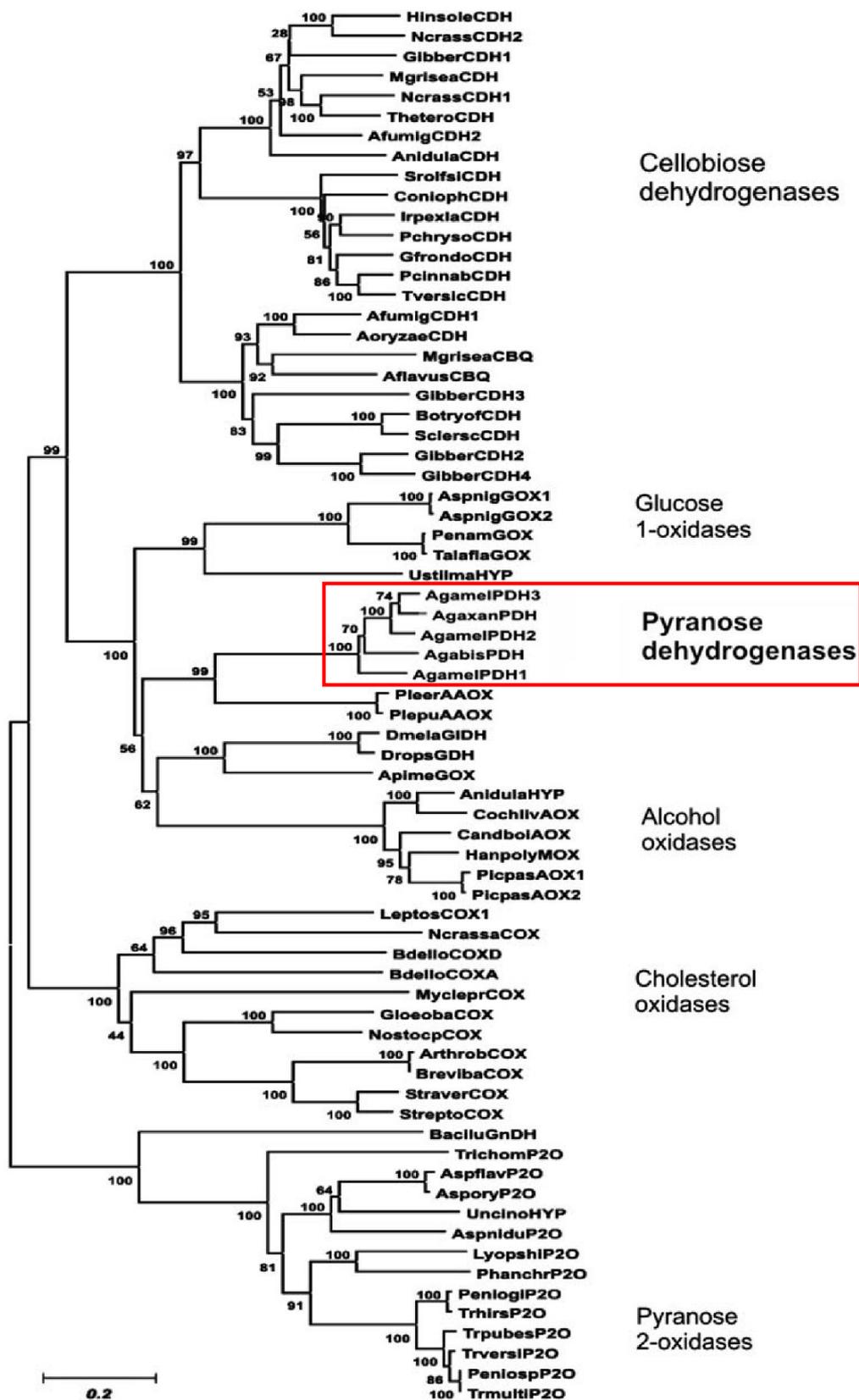


Figure 5: Reconstructed, unrooted phylogenetic tree of the GMC superfamily. Numbers on the branches represent bootstrap values obtained from 1,000 replicas. The scale bar represents 20% of the estimated sequence divergence. Pyranose dehydrogenase genes are in bold and encircled red. From Kittl and coworkers [28].

1.2.1 Pyranose dehydrogenase (PDH)

PDH is a monomer of approximately 65 kDa, carries a FAD cofactor in the active site, and is found in restricted groups of lignocellulolytic, litter-degrading basidiomycetes and some ascomycetes/fungi imperfecti [31]. It (di)oxidizes free, nonphosphorylated mono- and polysaccharides at C1–C4 (**Figure 6**), making it a very versatile catalyst for organic synthesis and bioelectrochemistry. PDH does not utilize oxygen as electron acceptor, but organometallic compounds or substituted quinones instead. Its biological role is not fully understood to date, but owing to its biochemical properties and its occurrence in litter decomposing fungi, it is probably involved in lignin degradation. Another possible role is the degradation of quinones, which some plants produce to defend themselves against attacks from fungi.

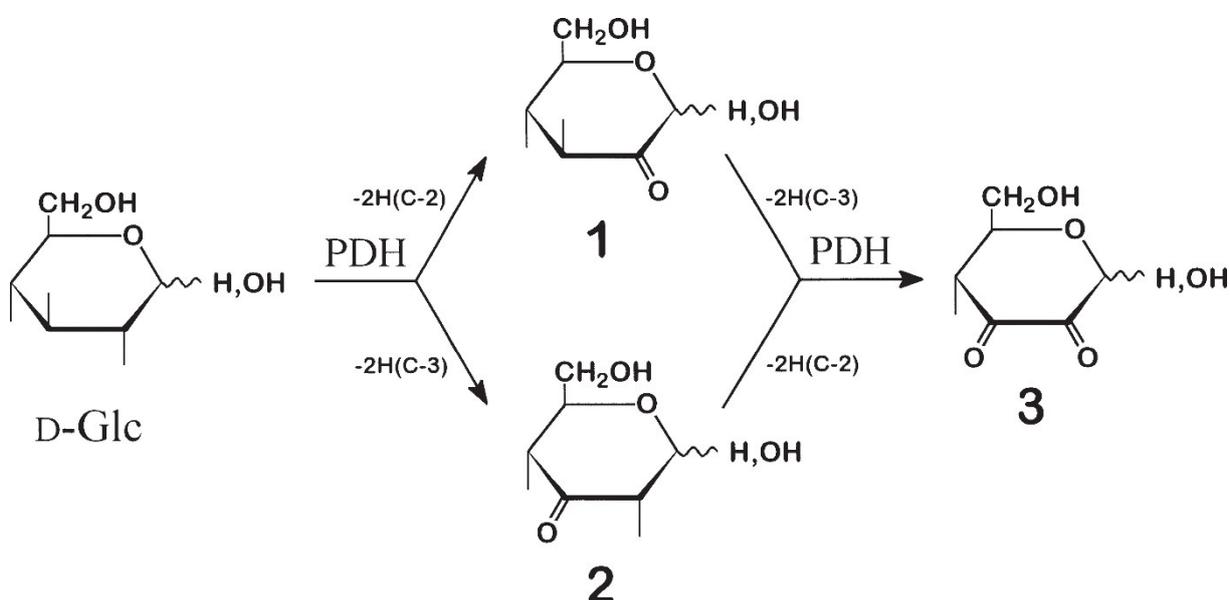


Figure 6: Proposed scheme for the dioxidation of D-glucose to the final product 2,3-diketo-D-glucose (3) by PDH from *Agaricus bisporus* (AbPDH). According to this scheme, AbPDH equally prefers both intermediates, 2-keto-D-glucose (1) and 3-keto-D-glucose (2). From Volc and coworkers [32].

PDH was first isolated from the fungus *Agaricus bisporus* [33] and *Macrolepiota rhacodes* [34], and only then from *Agaricus meleagris* (*Am*PDH) [35] and other sources such as *Agaricus xanthoderma* [36]. Aryl alcohol oxidases from the genus *Pleurotus* are the closest phylogenetic neighbors of PDH (**Figure 5**). In *Agaricus meleagris*, three genes for PDH isoforms with similarities of 80% or higher were identified (*pdh1*, *pdh2*, and *pdh3*) [28], but only the enzyme PDH1 was successfully isolated from the natural source as well a recombinantly produced and characterized to date. The heterologous expression and subsequent characterization of isoforms PDH2 and PDH3 is a matter of ongoing studies. The focus of this PhD thesis was to investigate *Am*PDH1, which is denoted *Am*PDH, PDH1, or simply PDH throughout this thesis, unless stated otherwise.

In the last years, *Am*PDH has received considerable attention. The wild-type enzyme from the native source *A. meleagris* was characterized in terms of its steady-state parameters for different substrates, its stability, and other properties [37]. However, the production of *Am*PDH in the native source is very time consuming and takes approximately six weeks to yield only a few mg of pure enzyme. Therefore, its heterologous expression was investigated, which resulted in the identification of *Pichia pastoris* as suitable expression system as indicated by high yields of soluble enzyme with similar kinetic properties compared to *Am*PDH from the native source [38]. This paved the way to provide sufficient amounts of *Am*PDH to elucidate its high resolution crystal structure [8] as depicted in **Figure 7A** (PDB code 4H7U). **Figure 7B** shows the most important active site residues. The crystal structure clearly indicates that an FAD is monocovalently attached to the N ϵ of His-103 of *Am*PDH via a 8 α -N $^{\beta}$ -histidyl-FAD linkage, which is the most common covalent FAD-linkage type [17]. *Am*PDH has two active site histidines (His-512 and His-556), which have been observed previously for the GMC members AAO and GOX. Other GMC family members such as CHO and P2O possess a catalytic His/Asn pair, where the histidine corresponds to His-512 in *Am*PDH and the asparagine to His-556 in *Am*PDH. The aim of this PhD thesis was to comprehensively study the role of these active site residues in *Am*PDH (**Figure 7B**).

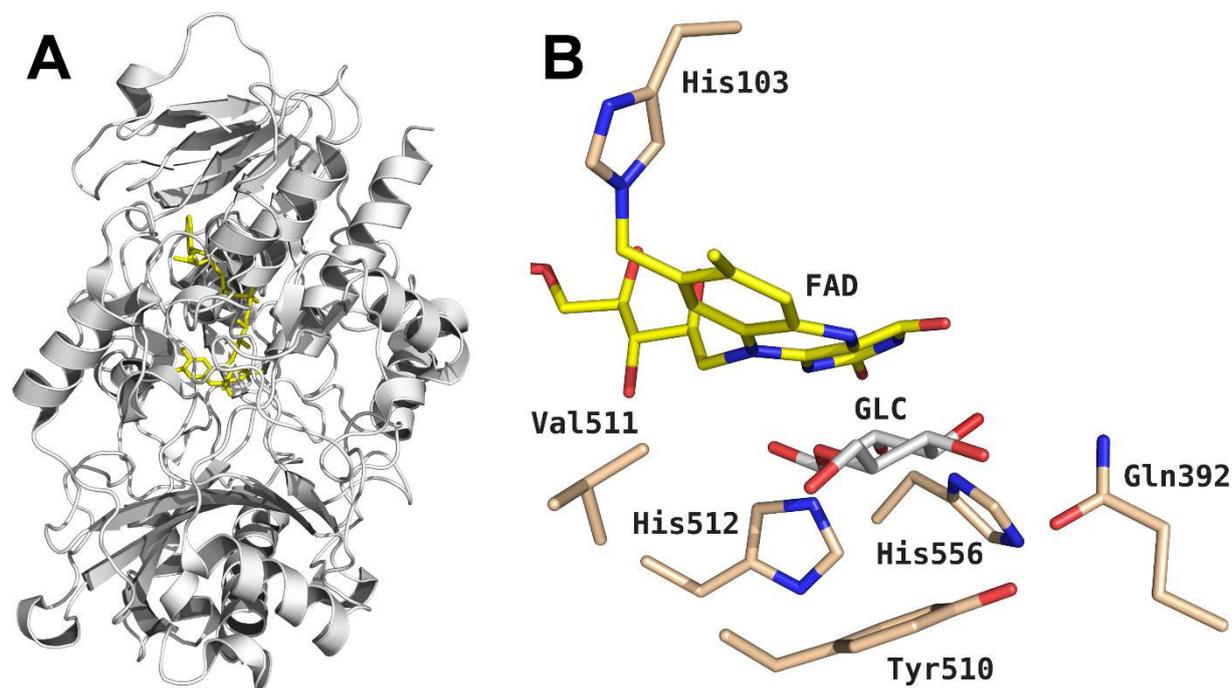


Figure 7: (A) 1.6 Å Crystal structure (PDB code 4H7U) of *Am*PDH elucidated by Tan and coworkers in 2013 [8]. (B) Active site residues of *Am*PDH that were investigated in the current thesis.

With the crystal structure and means to efficiently produce *Am*PDH at hand, subsequent mutational studies were readily possible. In a first attempt, variants with increased oxygen reactivity were identified by semi-rational protein engineering and subsequent high-throughput screening

[39]. Since it is still elusive why flavin-dependent enzymes react with oxygen or not, studies on the oxygen reactivity of flavoenzymes are of significant academic interest [40–42]. Moreover, large-scale applications of *Am*PDH require cheap electron acceptors such as O₂, because the electron acceptors described earlier are expensive and have to be efficiently removed for the desired applications. Krondorfer and coworkers recently identified one variant (H103Y) with about 5-fold increased oxygen reactivity compared to the wild-type [39]. In H103Y, the covalent linkage to the FAD *via* His-103 was disrupted but FAD was still tightly bound. In a subsequent study on this variant it was elucidated that its redox potential was decreased by about 100 mV compared to the wild type [9]. EPR spectroscopy revealed that the FAD-mobility within the active site was significantly increased and *Thermo*FAD experiments [43] demonstrated that the thermal stability of H103Y decreased by 8.2 °C when compared to the wild-type. This indicated the importance of the covalent FAD linkage for the studied properties.

Another research-area on PDH (especially *Am*PDH) is bioelectrochemistry [21,44,45]. Only recently, the ability of deglycosylated *Am*PDH to perform direct electron transfer (DET) to various electrodes was discovered [46–48]. DET is not dependent on mediators that shuttle electrons from the enzyme to the electrode, thereby interference from these mediators can be excluded [49]. This discovery could boost efforts to further engineer *Am*PDH towards bioelectrochemical applications.

1.2.2 Aryl-alcohol oxidase (AAO)

AAO is a monomeric oxidoreductase of about 62 kDa, carrying a non-covalently attached FAD cofactor [50]. It is found in *Pleurotus eryngii*, a basidiomycete belonging to the white-rot fungi, which is involved in lignin degradation. The ligninolytic system of *P. eryngii* is able to selectively degrade lignin but not cellulose, which makes it interesting for biotechnological applications [51]. An important part of this ligninolytic system is AAO, which is able to oxidize aromatic and aliphatic polyunsaturated, primary alcohols to their corresponding aldehydes [52]. The enzyme utilizes O₂ as electron acceptor and ultimately generates H₂O₂, which provides a continuous source of oxidative power for subsequent reactions by ligninolytic peroxidases. The active site in AAO contains a catalytic histidine pair His-502 (equivalent to His-512 in PDH) and His-546 (equivalent to His-556 in PDH) [8,53].

1.2.3 Choline oxidase (CHO)

CHO is a homodimer of about 120 kDa that carries a monocovalently attached FAD [54]. It catalyzes the four-electron oxidation of choline to glycine betaine *via* betaine aldehyde as intermediate and the concomitant reduction of molecular oxygen to H₂O₂ [55]. Glycine betaine accumulates to high levels in the cytoplasm and prevents dehydration and plasmolysis in

hyperosmotic environments [56]. This process is coupled to other physiological responses, *e.g.* regulation of internal ionic strength and pH as well as heat and cold tolerance [55]. Moreover, hyperosmolarity is associated with human pathogenicity [57]. Consequently, a mechanistic understanding of CHO and its adaption by enzyme engineering techniques has a potential impact in green and red biotechnology. CHO from *Arthrobacter globiformis* is readily available, as it can be heterologously expressed and purified in large amounts [55]. Extensive studies on the reaction mechanism of CHO were performed [58–60] and its oxygen reactivity reviewed recently [61]. The active site of CHO contains the catalytic pair His-466 (corresponding to His-512 in *Am*PDH) and Asn-510 (corresponding to His-556 in *Am*PDH) [62].

1.2.4 Glucose 1-oxidase (GOX)

GOX was discovered in *Aspergillus niger* extracts by Müller already in 1928 [63]. It is a homodimer with a molecular weight of about 150 kDa, whereas each subunit carries a tightly, but non-covalently bound FAD cofactor [64]. GOX oxidizes D-glucose at position C1 to δ -gluconolactone, which ultimately yields gluconic acid, and concomitantly reduces molecular oxygen to H_2O_2 [49]. Its main natural function is to generate H_2O_2 , which is targeted against bacteria and fungi lacking H_2O_2 scavengers such as catalase. GOX is also involved in lignin degradation, lowering the pH of the environment, and assisting in plant infection [49]. Industrial applications of GOX are manifold, making it somewhat a ‘star’ amongst commercially utilized enzymes. It is used as additive in food processing, for gluconic acid production, laundry detergents, textile bleaching, in immunoassays, and many other applications. The widest known field of GOX application is probably bioelectrochemistry, where it can be used in biofuel cells and – even more prominently – in glucose biosensors, where it can reliably determine blood-glucose levels. Especially the latter area is dominated by GOX: Newman and Turner recently estimated the value of world market for biosensors to be around 5 billion \$, whereas 85% were attributed to glucose biosensors [65]. Many of those glucose sensors are GOX-based. It is worthwhile mentioning that studies on the direct electron transfer (DET) of GOX immobilized on carbon nanotubes is a matter of ongoing studies and debate [66]. Similarly to AAO, GOX contains a catalytic histidine pair, consisting of His-516 (equivalent to His-512 in PDH) and His-559 (equivalent to His-556 in PDH), where His-516 was identified as the catalytic base in both, the reductive- and oxidative half reaction [64].

1.2.5 Pyranose 2-oxidase (P2O)

P2O is a homotetrameric enzyme with a molecular mass of about 270 kDa, in which each subunit carries a monocovalently attached FAD cofactor [67]. It is found in lignocellulolytic basidiomycetes and oxidizes preferentially D-glucose at position C2 to the corresponding 2-keto-D-glucose. Concomitantly, P2O reduces molecular oxygen to H_2O_2 , which is utilized by the fungi as a source

of oxidative power for their ligninolytic machinery involving peroxidases. P2O is interesting for industrial applications in biofuel cells [21], analytics [68], and the production of D-tagatose, a low-caloric, non-cariogenic sweetener [69]. P2O of *Trametes multicolor* (*Tm*P2O) can be easily heterologously expressed in *E. coli* [70] and has been investigated extensively – for a recent review, see [30]. High-resolution crystal structures of the wild-type enzyme [71] as well as of variants with substrates bound in different poses [72,73] are readily available. Similar to CHO, *Tm*P2O has a catalytic His/Asn pair in the active site, consisting of His-548 (corresponding to His-512 in *Am*PDH) and Asn-593 (corresponding to His-556 in *Am*PDH). Most recently, studies on the oxygen reactivity of active site variants [39] and on rationalizing the binding modes of fluorinated substrates [74] were performed.

1.3 Computational enzyme engineering and case studies on GMC oxidoreductases

Enzymes isolated from their natural source are rarely ideally adapted for applications in medicine and diagnostics as well as in environmental- or industrial biotechnology [75]. With the advent of molecular biology techniques, enzyme engineering has rapidly evolved as a standard technique to generate tailor-made biocatalysts for the desired applications [76]. One can distinguish between experimental and computational methods in enzyme engineering, which are ideally combined [77].

In experimental enzyme engineering, rational and random approaches can be distinguished. The rational engineering of a protein towards a desired function usually requires the availability of its structure as well as detailed knowledge about the mechanism as well as a functional assignment of key residues. Random approaches are more flexible in this respect, but require high-throughput screening of large combinatorial libraries, which can be very cumbersome and costly. Consequently, homology models of the protein of interest greatly facilitate engineering efforts – also in a random approach [78]. Experimental enzyme engineering techniques have already produced many artificial enzymes adapted for specific tasks. An example is the redesign of whole enzymes by iterative rounds of directed evolution to increase stability and tolerance to non-natural conditions, broaden substrate specificity and stereoselectivity, and enhance trace activities [75]. Another example is the generation of catalytic antibodies, which exploits the diversity and specificity of the mammalian immune system [79,80]. Haptens, which mimic the transition state of a chemical reaction, are used for immunization to generate catalytic antibodies with similar features to enzymes: substrate specificity, regio- and stereoselectivity, as well as significant rate accelerations compared to the uncatalyzed reaction. Nevertheless, the efficiencies of the best catalytic antibodies are usually orders of magnitudes lower compared to enzyme-catalyzed reactions, making catalysis *via* antibodies more primitive catalysts than enzymes [79]. This can largely be attributed to the fact that antibodies evolved to recognize and tightly bind one specific molecule, whereas enzymes show a higher

conformational flexibility, which contributes to efficient stabilization of reaction-intermediates and to exclude water from the active site [75].

If experimental engineering strategies have already yielded so many enzymes with the desired properties, why should one introduce a computational step? Computational considerations can substantially increase the frequency of finding active variants [81–83] and proteins with improved stability, new enzymatic activities, or new binding affinities [77]. Therefore, it seems worthwhile to consider computational enzyme engineering approaches, which will be introduced in the following section in more detail. The applied computational tools can be grouped in three categories: bioinformatics, molecular modeling, and *de novo* design [84].

1.3.1 Bioinformatics

Bioinformatics is an integrated approach to study and process biological data by computational means. Examples are bioinformatical web servers such as *ZEBRA*, which can systematically identify and analyze adaptive mutations in enzyme functional subfamilies [85]. Subfamily-specific positions (SSP) are conserved within these subfamilies but should differ among different subfamilies. Detailed statistical analysis of the SSPs is then translated into their significance. Based on this information, the SSP can subsequently be adapted by (semi)rational protein engineering. Another example for utilizing bioinformatics in enzyme engineering is the *JANUS* method, which predicts mutations to interconvert structurally related but functionally distinct enzymes by analyzing multiple-sequence alignments [86]. Another approach is to utilize the *Adaptive Substituent Reordering Algorithm* (ASRA) [87], a predictive bioinformatic model, which is trained on experimental data (**Figure 8**). ASRA follows a similar strategy as the well-established ‘quantitative structure–activity relationships’ (QSAR) methodology [81] and can readily be utilized in the directed evolution of enzymes. It detects the fundamental regularity of the protein property landscape and predicts properties of uncharacterized variants.

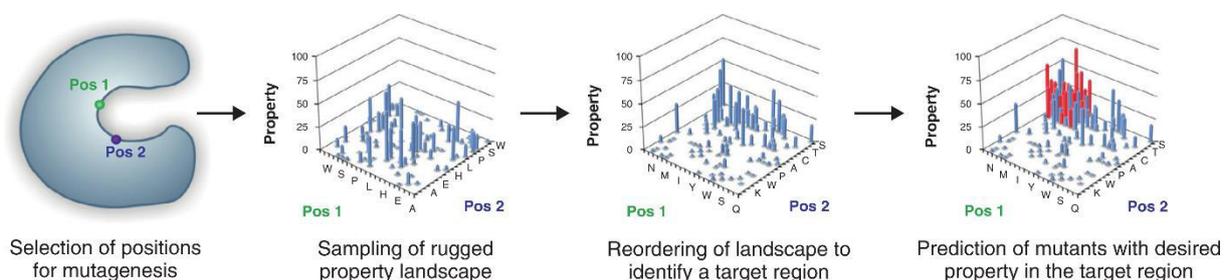


Figure 8: Workflow for Adaptive Substituent Reordering Algorithm (ASRA) to identify variants with desired properties. From Damborsky and Brezovsky [84].

1.3.2 *De novo* protein design

De novo enzyme design utilizes a known transition state for a targeted chemical reaction, around which an active site and eventually a whole enzyme is sculpted (**Figure 9A**) [88]. After identifying the rate-limiting transition state and calculating its structure by means of quantum mechanics, catalytic groups are determined *in silico* to stabilize it. The resulting ensemble is an idealized, 3D-model of a minimal active site, termed theoretical enzyme (“theozyme”) [89]. This theozyme is docked *in silico* into protein structures from the PDB with programs such as RosettaMatch [90], Scaffold Select [91], and ORBIT [92]. Residues in and around this pocket are then mutated for optimal packing of the transition state and catalytic groups. Each possible sequence is then energy-minimized by iteratively searching the conformational space of the ligand and the side chains. Finally, the most promising designs are ranked according to their calculated energies and tested *in vitro*. This process can be compared to antibody affinity maturation in the immune system, with the stability of the fold serving as the selection criterion for passing a design on to the next round of optimization [75]. *De novo* designed enzymes of the first-generation usually have low catalytic efficiencies, which can be significantly increased by subsequent rounds of redesigning, which are guided by biochemical information from earlier rounds, or by directed evolution [75,88,93]. Therefore, computational enzyme engineering and experimental engineering approaches are complementing strategies to obtain tailor-made biocatalysts [94]. **Figure 9B** gives an example of computational *de novo* design and directed evolution of a Kemp eliminase [95].

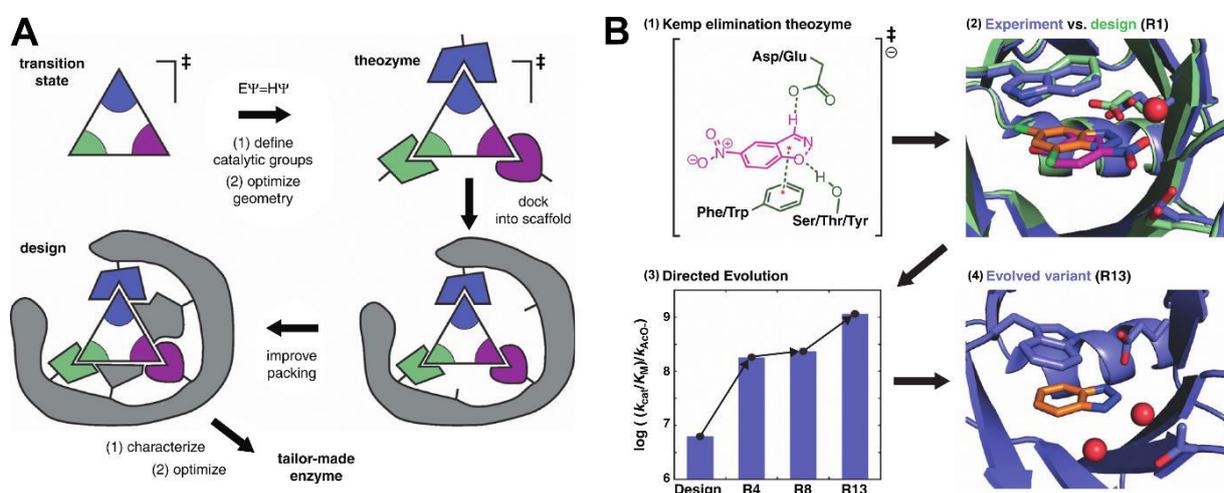


Figure 9: (A) Flow-scheme for the computational *de novo* design methodology. See text for a detailed description (B) Flow-scheme for the computational *de novo* design in combination with directed evolution of a Kemp eliminase. The theozyme (B1) was used as a template for subsequent design. (B2) The crystal structure of the first-generation enzyme (blue) is in good qualitative agreement with the computationally designed model (green). The ligand from the crystal structure (orange) is flipped relative to the designed orientation (pink), though. (B3) Catalytic efficiencies steadily increased from the first round (R2) until the last round of mutagenesis and screening (R13) relative to the designed starting enzyme (R1; Design). (B4) The evolved variant (R13) shows structural adaptations in the active site, which are responsible for its higher catalytic efficiency. From Kries and coworkers [88].

1.3.3 Molecular modeling

From an enzyme engineering perspective, various approaches can be assigned to molecular modeling techniques [84]. Examples are finding a suitable ligand for an enzyme *via* docking and scoring or to probe its dynamic properties *via* molecular dynamics (MD) and quantum mechanics (QM) simulations. The trajectories of MD simulations can subsequently be analyzed for different properties *e.g.* the accessibility of the active site through substrate channels with the software tool *CAVER* [96] or relative binding free energies of different compounds to the active site, which yields the best binding substance for a given enzyme [97]. Because MD simulations do not take into account particles that are smaller than whole atoms, QM simulations have to be conducted in order to obtain information about sub-atomic processes [98]. However, QM simulations are computationally very expensive. To address this problem, hybrid approaches were developed, that use a QM method for the chemically active region of an enzyme and combine it with a molecular mechanics (MM) treatment for the rest of the enzyme (QM/MM simulations) [99,100]. Consequently, the level of detail for modeling the enzyme of interest is dependent on the type of process one is interested. In the course of this PhD thesis, biomolecular properties of *AmPDH* were investigated that were readily accessible *via* MD simulations. **Figure 10A** illustrates four biomolecular processes that are governed by thermodynamic equilibria that play a fundamental role in the behavior of these systems. These four processes are all driven by weak, nonbonded interatomic interactions, which determine the thermodynamic properties of the condensed phase in which the four processes occur. Consequently, these processes are ideally modeled at the atomic or molecular level [98]. Lastly, the four basic choices when modeling a biomolecular system are shown in **Figure 10B**.

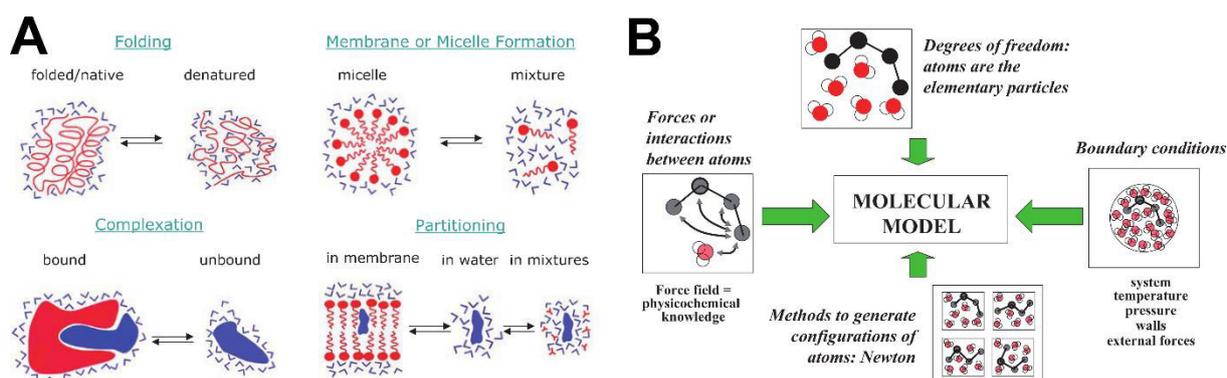


Figure 10: (A) Four biomolecular processes that are governed by thermodynamic equilibria. (B) Four basic choices when defining a model for molecular simulation. From van Gunsteren and coworkers [98].

1.3.4 Case studies on computational investigations of GMC oxidoreductases

In this section, short examples for computational studies on GMC oxidoreductases AAO, CHO, GOX, and P2O are given.

For AAO, substrate diffusion, oxidation, and stereoselectivity as well as the role of the active site histidines in catalysis was investigated by Hernández-Ortega and coworkers by combining computational and experimental efforts [52,101,102]. The alcohol-migration pathway into the active site and the protonation states of the two active site histidines were determined by computational means. Moreover, the timing of proton and hydride transfers and the existence of reaction intermediates were probed by computing energy profiles of the AAO-catalyzed reaction by QM/MM methods [101]. QM/MM calculations corroborated the experimental observation that only active site His-502 is involved in the oxidative half reaction of AAO with O₂ [52]. Lastly, the stereoselectivity of AAO towards *p*-methoxybenzyl alcohol could be attributed to the arrangement of the hydride-receiving atoms (flavin N5 and His-502 N ϵ) in relation to the alcohol C α -substituents *via* substrate-docking into the active site and subsequent QM/MM calculations [102].

For CHO, a detailed computational study on the active site gating revealed that not a presumptive active site loop controls its access but a patch of hydrophobic residues, which open and close on a picosecond time scale [103]. Interestingly, the authors suggest that a similar gating mechanism could play a role in other GMC oxidoreductases and that one should be careful when claiming that a loop controls access to the active site. In the same study, electrostatic interactions were identified to govern the rate constant for formation of the choline-CHO complex.

For GOX, docking experiments and subsequent MD simulations were applied to determine the most likely orientation of substrates and products in the active site [64,104]. This corroborated the proposed reaction mechanism obtained from experiments [64]. QM/MM calculations were used to determine the role of the FAD and active site residues in D-glucose oxidation [104]. With this approach, heats of formation for the enzyme–substrate or enzyme–product complexes were computed. The active site histidine His-516 was proposed to be involved in substrate binding and C4a-hydroperoxyflavin stabilization, whereas His-559 was suggested to act as proton acceptor.

For P2O, the physical factors in the oxidative half-reaction of reduced flavin with O₂ could be identified by combining QM analysis and transient kinetic studies, which rationalized oxygen activation and C4a-hydroperoxyflavin stabilization [105]. In another recent study on P2O, MD simulations were used to investigate the structural heterogeneity of the four subunits and how this heterogeneity affects ultrafast fluorescence dynamics [106]. Related to the O₂ reactivity of P2O, general principles of O₂ diffusion into flavoenzymes were investigated in a study combining MD simulations with experiments [107]. This study indicated that multiple pathways effectively guide O₂ diffusion to a few key residues or atoms in the active site, *e.g.* the flavin C4a atom. Nevertheless, the microenvironment of the flavin was found to be very important for modulating the O₂ reactivity to yield dehydrogenases, monooxygenases, or oxidases [42].

1.4 References

- 1 Joosten V & Van Berkel W (2007) Flavoenzymes. *Curr. Opin. Chem. Biol.* **11**, 195–202.
- 2 Macheroux P (1999) UV-visible spectroscopy as a tool to study flavoproteins. *Flavoprotein protocols (Chapman SK & Reid GA Eds.)*. Springer Science & Business Media. pp. 1–7.
- 3 Aliverti A, Curti B & Vanoni MA (1999) Identifying and Quantitating FAD and FMN in Simple and in Iron-Sulfur-Containing Flavoproteins. *Flavoprotein protocols (Chapman SK & Reid GA eds.)*. Springer Science & Business Media. pp. 9–23.
- 4 Koziol J (1971) Fluorometric analyses of riboflavin and its coenzymes. *Methods in Enzymology (Donald B. McCormick & Lemuel D. Wright, eds)*. Springer Science & Business Media. pp. 253–285.
- 5 Hinkson JW (1968) Azotobacter free-radical flavoprotein. Preparation and properties of the apoprotein. *Biochemistry* **7**, 2666–2672.
- 6 Macheroux P, Kappes B & Ealick SE (2011) Flavogenomics - a genomic and structural view of flavin-dependent proteins: Flavin-dependent proteins. *FEBS J.* **278**, 2625–2634.
- 7 Walsh CT & Wencewicz TA (2013) Flavoenzymes: Versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **30**, 175.
- 8 Tan TC, Spadiut O, Wongnate T, Sucharitakul J, Krondorfer I, Sygmund C, Haltrich D, Chaiyen P, Peterbauer CK & Divne C (2013) The 1.6 Å crystal structure of pyranose dehydrogenase from *Agaricus meleagris* rationalizes substrate specificity and reveals a flavin ontermediate. *PLoS ONE* **8**, e53567.
- 9 Krondorfer I, Brugger D, Paukner R, Scheiblbrandner S, Pirker KF, Hofbauer S, Furtmüller PG, Obinger C, Haltrich D & Peterbauer CK (2014) *Agaricus meleagris* pyranose dehydrogenase: Influence of covalent FAD linkage on catalysis and stability. *Arch. Biochem. Biophys.* **558**, 111–119.
- 10 Massey V (1995) Introduction: Flavoprotein structure and mechanism. *FASEB J.* **9**, 473–475.
- 11 Fraaije MW & Mattevi A (2000) Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci.* **25**, 126–132.
- 12 Ghisla S & Massey V (1989) Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.* **181**, 1–17.
- 13 Fitzpatrick PF (2001) Substrate dehydrogenation by flavoproteins. *Acc. Chem. Res.* **34**, 299–307.
- 14 Porter DJT, Voet JG & Bright HJ (1973) Direct evidence for carbanions and covalent N5-flavin-carbanion adducts as catalytic intermediates in the oxidation of nitroethane by D-amino acid oxidase. *J. Biol. Chem.* **248**, 4400–4416.
- 15 Thorpe C & Williams CH (1976) Spectral evidence for a flavin adduct in a monoalkylated derivative of pig heart lipoamide dehydrogenase. *J. Biol. Chem.* **251**, 7726–7728.
- 16 Massey V (1994) Activation of molecular oxygen by flavins and flavoproteins. *J. Biol. Chem.* **269**, 22459–22462.
- 17 Heuts DPHM, Scrutton NS, McIntire WS & Fraaije MW (2009) What's in a covalent bond? On the role and formation of covalently bound flavin cofactors. *FEBS J.* **276**, 3405–3427.

- 18 Koetter JWA & Schulz GE (2005) Crystal structure of 6-hydroxy-D-nicotine oxidase from *Arthrobacter nicotinovorans*. *J. Mol. Biol.* **352**, 418–428.
- 19 Hassan-Abdallah A, Bruckner RC, Zhao G & Jorns MS (2005) Biosynthesis of covalently bound flavin: Isolation and *in vitro* flavinylation of monomeric sarcosine oxidase apoprotein. *Biochemistry* **44**, 6452–6462.
- 20 May SW & Padgett SR (1983) Oxidoreductase enzymes in biotechnology: Current status and future potential. *Nat. Biotechnol.* **1**, 677–686.
- 21 Tasca F, Timur S, Ludwig R, Haltrich D, Volc J, Antiochia R & Gorton L (2007) Amperometric biosensors for detection of sugars based on the electrical wiring of different pyranose oxidases and pyranose dehydrogenases with osmium redox polymer on graphite electrodes. *Electroanalysis* **19**, 294–302.
- 22 Cracknell JA, Vincent KA & Armstrong FA (2008) Enzymes as working or inspirational electrocatalysts for fuel cells and electrolysis. *Chem. Rev.* **108**, 2439–2461.
- 23 Grieshaber D, MacKenzie R, Voeroes J & Reimhult E (2008) Electrochemical biosensors—Sensor principles and architectures. *Sensors* **8**, 1400–1458.
- 24 Cavener DR (1992) GMC oxidoreductases: A newly defined family of homologous proteins with diverse catalytic activities. *J. Mol. Biol.* **223**, 811–814.
- 25 Mattevi A (1998) The PHBH fold: Not only flavoenzymes. *Biophys. Chem.* **70**, 217–222.
- 26 Kiess M, Hecht H-J & Kalisz HM (1998) Glucose oxidase from *Penicillium amagasakiense*. *Eur. J. Biochem.* **252**, 90–99.
- 27 Martin Hallberg B, Henriksson G, Pettersson G & Divne C (2002) Crystal structure of the flavoprotein domain of the extracellular flavocytochrome cellobiose dehydrogenase. *J. Mol. Biol.* **315**, 421–434.
- 28 Kittl R, Sygmund C, Halada P, Volc J, Divne C, Haltrich D & Peterbauer CK (2008) Molecular cloning of three pyranose dehydrogenase-encoding genes from *Agaricus meleagris* and analysis of their expression by real-time RT-PCR. *Curr. Genet.* **53**, 117–127.
- 29 Zámocký M, Hallberg M, Ludwig R, Divne C & Haltrich D (2004) Ancestral gene fusion in cellobiose dehydrogenases reflects a specific evolution of GMC oxidoreductases in fungi. *Gene* **338**, 1–14.
- 30 Wongnate T & Chaiyen P (2013) The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose–methanol–choline superfamily. *FEBS J.* **280**, 3009–3027.
- 31 Peterbauer CK & Volc J (2010) Pyranose dehydrogenases: Biochemical features and perspectives of technological applications. *Appl. Microbiol. Biotechnol.* **85**, 837–848.
- 32 Volc J, Sedmera P, Halada P, Přikrylová V & Daniel G (1998) C-2 and C-3 oxidation of D-Glc, and C-2 oxidation of D-Gal by pyranose dehydrogenase from *Agaricus bisporus*. *Carbohydr. Res.* **310**, 151–156.
- 33 Volc J, Kubátová E, Wood DA & Daniel G (1997) Pyranose 2-dehydrogenase, a novel sugar oxidoreductase from the basidiomycete fungus *Agaricus bisporus*. *Arch. Microbiol.* **167**, 119–125.

- 34 Volc J, Kubátová E, Daniel G, Sedmera P & Haltrich D (2001) Screening of basidiomycete fungi for the quinone-dependent sugar C-2/C-3 oxidoreductase, pyranose dehydrogenase, and properties of the enzyme from *Macrolepiota rhacodes*. *Arch. Microbiol.* **176**, 178–186.
- 35 Volc J, Sedmera P, Halada P, Daniel G, Příkrylová V & Haltrich D (2002) C-3 oxidation of non-reducing sugars by a fungal pyranose dehydrogenase: Spectral characterization. *J. Mol. Catal. - B Enzym.* **17**, 91–100.
- 36 Volc J, Sedmera P, Kujawa M, Halada P, Kubátová E & Haltrich D (2004) Conversion of lactose to β -D-galactopyranosyl-(1 \rightarrow 4)-D-arabino-hexos-2-ulose-(2-dehydrolactose) and lactobiono-1,5-lactone by fungal pyranose dehydrogenase. *J. Mol. Catal. B Enzym.* **30**, 177–184.
- 37 Sygmund C, Kittl R, Volc J, Halada P, Kubátová E, Haltrich D & Peterbauer CK (2008) Characterization of pyranose dehydrogenase from *Agaricus meleagris* and its application in the C-2 specific conversion of D-galactose. *J. Biotechnol.* **133**, 334–342.
- 38 Sygmund C, Gutmann A, Krondorfer I, Kujawa M, Glieder A, Pscheidt B, Haltrich D, Peterbauer C & Kittl R (2012) Simple and efficient expression of *Agaricus meleagris* pyranose dehydrogenase in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **94**, 695–704.
- 39 Krondorfer I, Lipp K, Brugger D, Staudigl P, Sygmund C, Haltrich D & Peterbauer CK (2014) Engineering of pyranose dehydrogenase for increased oxygen reactivity. *PLoS ONE* **9**, e91145.
- 40 Mattevi A (2006) To be or not to be an oxidase: Challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* **31**, 276–283.
- 41 McDonald CA, Fagan RL, Collard F, Monnier VM & Palfey BA (2011) Oxygen reactivity in flavoenzymes: Context matters. *J. Am. Chem. Soc.* **133**, 16809–16811.
- 42 Chaiyen P, Fraaije MW & Mattevi A (2012) The enigmatic reaction of flavins with oxygen. *Trends Biochem. Sci.* **37**, 373–380.
- 43 Forneris F, Orru R, Bonivento D, Chiarelli LR & Mattevi A (2009) *ThermoFAD*, a *Thermofluor*-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* **276**, 2833–2840.
- 44 Tasca F, Gorton L, Kujawa M, Patel I, Harreither W, Peterbauer CK, Ludwig R & Nöll G (2010) Increasing the coulombic efficiency of glucose biofuel cell anodes by combination of redox enzymes. *Biosens. Bioelectron.* **25**, 1710–1716.
- 45 Zafar MN, Tasca F, Boland S, Kujawa M, Patel I, Peterbauer CK, Leech D & Gorton L (2010) Wiring of pyranose dehydrogenase with osmium polymers of different redox potentials. *Bioelectrochemistry* **80**, 38–42.
- 46 Yakovleva ME, Killyéni A, Ortiz R, Schulz C, MacAodha D, Conghaile PÓ, Leech D, Popescu IC, Gonaus C, Peterbauer CK & Gorton L (2012) Recombinant pyranose dehydrogenase—A versatile enzyme possessing both mediated and direct electron transfer. *Electrochem. Commun.* **24**, 120–122.
- 47 Yakovleva ME, Killyéni A, Seubert O, Ó Conghaile P, MacAodha D, Leech D, Gonaus C, Popescu IC, Peterbauer CK, Kjellström S & Gorton L (2013) Further insights into the catalytical properties of deglycosylated pyranose dehydrogenase from *Agaricus meleagris* recombinantly expressed in *Pichia pastoris*. *Anal. Chem.* **85**, 9852–9858.

- 48 Killyéni A, Yakovleva ME, MacAodha D, Conghaile PÓ, Gonaus C, Ortiz R, Leech D, Popescu IC, Peterbauer CK & Gorton L (2014) Effect of deglycosylation on the mediated electrocatalytic activity of recombinantly expressed *Agaricus meleagris* pyranose dehydrogenase wired by osmium redox polymer. *Electrochimica Acta* **126**, 61–67.
- 49 Wong CM, Wong KH & Chen XD (2008) Glucose oxidase: Natural occurrence, function, properties and industrial applications. *Appl. Microbiol. Biotechnol.* **78**, 927–938.
- 50 Ruiz-Dueñas FJ, Ferreira P, Martínez MJ & Martínez AT (2006) *In vitro* activation, purification, and characterization of *Escherichia coli* expressed aryl-alcohol oxidase, a unique H₂O₂-producing enzyme. *Protein Expr. Purif.* **45**, 191–199.
- 51 Valmaseda M, Almendros G & Martínez AT (1990) Substrate-dependent degradation patterns in the decay of wheat straw and beech wood by ligninolytic fungi. *Appl. Microbiol. Biotechnol.* **33**, 481–484.
- 52 Hernández-Ortega A, Lucas F, Ferreira P, Medina M, Guallar V & Martínez AT (2012) Role of active site histidines in the two half-reactions of the aryl-alcohol oxidase catalytic cycle. *Biochemistry* **51**, 6595–6608.
- 53 Hecht HJ, Kalisz HM, Hendle J, Schmid RD & Schomburg D (1993) Crystal structure of glucose oxidase from *Aspergillus niger* refined at 2.3 Å resolution. *J. Mol. Biol.* **229**, 153–172.
- 54 Rand T, Halkier T & Hansen OC (2003) Structural characterization and mapping of the covalently linked FAD cofactor in choline oxidase from *Arthrobacter globiformis*. *Biochemistry* **42**, 7188–7194.
- 55 Fan F, Ghanem M & Gadda G (2004) Cloning, sequence analysis, and purification of choline oxidase from *Arthrobacter globiformis*: A bacterial enzyme involved in osmotic stress tolerance. *Arch. Biochem. Biophys.* **421**, 149–158.
- 56 Burg MB, Kwon ED & Kültz D (1997) Regulation of gene expression by hypertonicity. *Annu. Rev. Physiol.* **59**, 437–455.
- 57 Badger JL & Kim KS (1998) Environmental growth conditions influence the ability of *Escherichia coli* K1 to invade brain microvascular endothelial cells and confer serum resistance. *Infect. Immun.* **66**, 5692–5697.
- 58 Fan & Gadda G (2005) On the catalytic mechanism of choline oxidase. *J. Am. Chem. Soc.* **127**, 2067–2074.
- 59 Gadda G (2008) Hydride transfer made easy in the reaction of alcohol oxidation catalyzed by flavin-dependent oxidases. *Biochemistry* **47**, 13745–13753.
- 60 Rungsririyachai K (2010) On the catalytic roles of HIS351, ASN510, and HIS466 in choline oxidase and the kinetic mechanism of pyranose 2-oxidase. *Dissertation, Georgia State University*.
- 61 Gadda G (2012) Oxygen activation in flavoprotein oxidases: the importance of being positive. *Biochemistry* **51**, 2662–2669.
- 62 Quaye O, Lountos GT, Fan, Orville AM & Gadda G (2007) Role of Glu312 in binding and positioning of the substrate for the hydride transfer reaction in choline oxidase. *Biochemistry* **47**, 243–256.

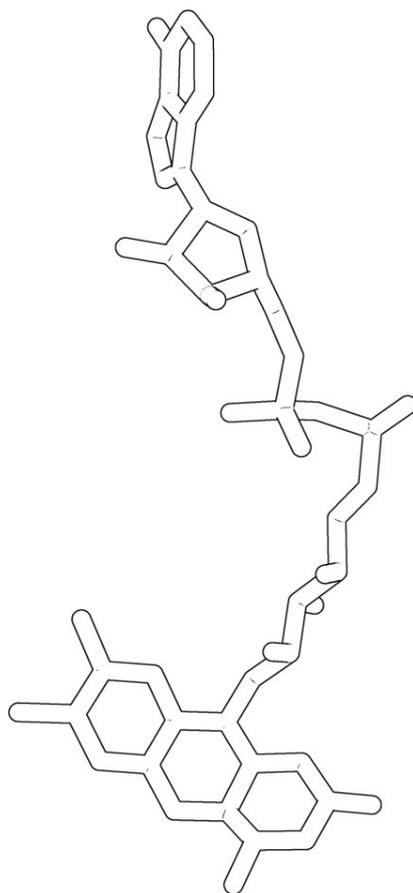
- 63 Müller D (1928) Oxidation von Glukose mit Extrakten aus *Aspergillus niger*. *Biochem. Z.* **199**, 136–170.
- 64 Wohlfahrt G, Trivić S, Zeremski J, Peričin D & Leskovac V (2004) The chemical mechanism of action of glucose oxidase from *Aspergillus niger*. *Mol. Cell. Biochem.* **260**, 69–83.
- 65 Newman JD & Turner APF (2005) Home blood glucose biosensors: A commercial perspective. *Biosens. Bioelectron.* **20**, 2435–2453.
- 66 Wooten M, Karra S, Zhang M & Gorski W (2014) On the direct electron transfer, sensing, and enzyme activity in the glucose oxidase/carbon nanotubes system. *Anal. Chem.* **86**, 752–757.
- 67 Leitner C, Volc J & Haltrich D (2001) Purification and characterization of pyranose oxidase from the white rot fungus *Trametes multicolor*. *Appl. Environ. Microbiol.* **67**, 3636–3644.
- 68 Giffhorn F (2000) Fungal pyranose oxidases: Occurrence, properties and biotechnical applications in carbohydrate chemistry. *Appl. Microbiol. Biotechnol.* **54**, 727–740.
- 69 Haltrich D, Leitner C, Neuhauser W, Nidetzky B, Kulbe KD & Volc J (1998) A convenient enzymatic procedure for the production of aldose-free D-tagatose. *Ann. N. Y. Acad. Sci.* **13**, 295–299.
- 70 Spadiut O, Posch G, Ludwig R, Haltrich D & Peterbauer CK (2010) Evaluation of different expression systems for the heterologous expression of pyranose 2-oxidase from *Trametes multicolor* in *E. coli*. *Microb. Cell Factories* **9**:14.
- 71 Martin Hallberg B, Leitner C, Haltrich D & Divne C (2004) Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase. *J. Mol. Biol.* **341**, 781–796.
- 72 Kujawa M, Ebner H, Leitner C, Hallberg BM, Prongjit M, Sucharitakul J, Ludwig R, Rudsander U, Peterbauer C, Chaiyen P, Haltrich D & Divne C (2006) Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J. Biol. Chem.* **281**, 35104–35115.
- 73 Tan T-C, Haltrich D & Divne C (2011) Regioselective control of β -D-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy. *J. Mol. Biol.* **409**, 588–600.
- 74 Tan TC, Spadiut O, Gandini R, Haltrich D & Divne C (2014) Structural basis for binding of fluorinated glucose and galactose to *Trametes multicolor* pyranose 2-oxidase variants with improved galactose conversion. *PLoS ONE* **9**, e86736.
- 75 Hilvert D (2013) Design of protein catalysts. *Annu. Rev. Biochem.* **82**, 447–470.
- 76 Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC & Robins K (2012) Engineering the third wave of biocatalysis. *Nature* **485**, 185–194.
- 77 Pantazes RJ, Grisewood MJ & Maranas CD (2011) Recent advances in computational protein design. *Curr. Opin. Struct. Biol.* **21**, 467–472.
- 78 Davids T, Schmidt M, Böttcher D & Bornscheuer UT (2013) Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr. Opin. Chem. Biol.* **17**, 215–220.
- 79 Hilvert D (2000) Critical analysis of antibody catalysis. *Annu. Rev. Biochem.* **69**, 751–793.

- 80 Schultz PG, Yin J & Lerner RA (2002) The chemistry of the antibody molecule. *Angew. Chem. Int. Ed.* **41**, 4427–4437.
- 81 Braiuca P, Ebert C, Basso A, Linda P & Gardossi L (2006) Computational methods to rationalize experimental strategies in biocatalysis. *Trends Biotechnol.* **24**, 419–425.
- 82 Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF, Hilvert D, Houk KN, Stoddard BL & Baker D (2008) *De novo* computational design of retro-aldol enzymes. *Science* **319**, 1387–1391.
- 83 Kipnis Y & Baker D (2012) Comparison of designed and randomly generated catalysts for simple chemical reactions. *Protein Sci.* **21**, 1388–1395.
- 84 Damborsky J & Brezovsky J (2014) Computational tools for designing and engineering enzymes. *Curr. Opin. Chem. Biol.* **19**, 8–16.
- 85 Suplatov D, Kirilin E, Takhaveev V & Švedas V (2013) Zebra: A web server for bioinformatic analysis of diverse protein families. *J. Biomol. Struct. Dyn.* **32**, 1752–1758.
- 86 Addington TA, Mertz RW, Siegel JB, Thompson JM, Fisher AJ, Filkov V, Fleischman NM, Suen AA, Zhang C & Toney MD (2013) Janus: Prediction and ranking of mutations required for functional interconversion of enzymes. *J. Mol. Biol.* **425**, 1378–1389.
- 87 Feng X, Sanchis J, Reetz MT & Rabitz H (2012) Enhancing the efficiency of directed evolution in focused enzyme libraries by the adaptive substituent reordering algorithm. *Chem. – Eur. J.* **18**, 5646–5654.
- 88 Kries H, Blomberg R & Hilvert D (2013) *De novo* enzymes by computational design. *Curr. Opin. Chem. Biol.* **17**, 221–228.
- 89 Tantillo DJ, Jiangang C & Houk KN (1998) Theozymes and compuzymes: Theoretical models for biological catalysis. *Curr. Opin. Chem. Biol.* **2**, 743–750.
- 90 Zanghellini A, Jiang L, Wollacott AM, Cheng G, Meiler J, Althoff EA, Röthlisberger D & Baker D (2006) New algorithms and an *in silico* benchmark for computational enzyme design. *Protein Sci.* **15**, 2785–2794.
- 91 Malisi C, Kohlbacher O & Höcker B (2009) Automated scaffold selection for enzyme design. *Proteins Struct. Funct. Bioinforma.* **77**, 74–83.
- 92 Bolon DN & Mayo SL (2001) Enzyme-like proteins by computational design. *Proc. Natl. Acad. Sci.* **98**, 14274–14279.
- 93 Wijma HJ & Janssen DB (2013) Computational design gains momentum in enzyme catalysis engineering. *FEBS J.* **280**, 2948–2960.
- 94 Kiss G, Çelebi-Ölçüm N, Moretti R, Baker D & Houk KN (2013) Computational enzyme design. *Angew. Chem. Int. Ed.* **52**, 5700–5725.
- 95 Khersonsky O, Kiss G, Röthlisberger D, Dym O, Albeck S, Houk KN, Baker D & Tawfik DS (2012) Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59. *Proc. Natl. Acad. Sci.* **109**, 10358–10363.
- 96 Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L, Sochor J & Damborsky J (2012) CAVER 3.0: A tool

- for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* **8**, e1002708.
- 97 Oostenbrink C (2012) Free energy calculations from one-step perturbations. *Computational drug discovery and design* (Baron, R., Ed). Humana Press, New York. Vol. 819, pp 487–499.
- 98 Van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, Gee P, Geerke DP, Glättli A, Hünenberger PH, Kastenholz MA, Oostenbrink C, Schenk M, Trzesniak D, van der Vegt NFA & Yu HB (2006) Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Ed Engl.* **45**, 4064–4092.
- 99 Senn HM & Thiel W (2007) QM/MM studies of enzymes. *Curr. Opin. Chem. Biol.* **11**, 182–187.
- 100 Meier K, Choutko A, Dolenc J, Eichenberger AP, Riniker S & van Gunsteren WF (2013) Multi-resolution simulation of biomolecular systems: A review of methodological issues. *Angew. Chem. Int. Ed.* **52**, 2820–2834.
- 101 Hernández-Ortega A, Borrelli K, Ferreira P, Medina M, Martínez AT & Guallar V (2011) Substrate diffusion and oxidation in GMC oxidoreductases: An experimental and computational study on fungal aryl-alcohol oxidase. *Biochem. J.* **436**, 341–350.
- 102 Hernández-Ortega A, Ferreira P, Merino P, Medina M, Guallar V & Martínez AT (2012) Stereoselective hydride transfer by aryl-alcohol oxidase, a member of the GMC superfamily. *ChemBioChem* **13**, 427–435.
- 103 Xin Y, Gadda G & Hamelberg D (2009) The cluster of hydrophobic residues controls the entrance to the active site of choline oxidase. *Biochemistry* **48**, 9599–9605.
- 104 Meyer M, Wohlfahrt G, Knäblein J & Schomburg D (1998) Aspects of the mechanism of catalysis of glucose oxidase: A docking, molecular mechanics and quantum chemical study. *J. Comput. Aided Mol. Des.* **12**, 425–440.
- 105 Wongnate T, Surawatanawong P, Visitsatthawong S, Sucharitakul J, Scrutton NS & Chaiyen P (2014) Proton-coupled electron transfer and adduct configuration are important for C4a-hydroperoxyflavin formation and stabilization in a flavoenzyme. *J. Am. Chem. Soc.* **136**, 241–253.
- 106 Lugsanangarm K, Kokpol S, Nueangaudom A, Pianwanit S, Nunthaboot N & Tanaka F (2014) Structural heterogeneity among four subunits in pyranose 2-oxidase: A molecular dynamics simulation study. *J. Theor. Comput. Chem.* **13**, 1440010.
- 107 Baron R, Riley C, Chenprakhon P, Thotsaporn K, Winter RT, Alfieri A, Forneris F, Van Berkel WJH, Chaiyen P, Fraaije MW, Mattevi A & McCammon JA (2009) Multiple pathways guide oxygen diffusion into flavoenzyme active sites. *Proc. Natl. Acad. Sci.* **106**, 10603–10608.

CHAPTER 2

Aim and Outline



Author

Michael Graf

2.1 Aim and outline of the thesis

Within the context of the considerations and methods mentioned in the previous section, the aim of this PhD thesis was to characterize the active site of the GMC oxidoreductase pyranose dehydrogenase from the fungus *Agaricus meleagris* (*Am*PDH). Special emphasis was placed on the covalent attachment of the FAD, the vicinity of the isoalloxazine, and the interactions of active site residues with substrates. This contributed to a better appreciation of fundamental biochemical and biophysical aspects of key residues in *Am*PDH, which may subsequently lead to enhanced protein engineering strategies to better adapt *Am*PDH towards the desired applications.

To achieve that goal, a combined experimental and computational approach was utilized:

- 1.) Active site variants of *Am*PDH were generated, heterologously expressed in *Pichia pastoris*, and subsequently characterized by biochemical and biophysical means.
- 2.) Molecular dynamics (MD) simulations and free energy calculations of *Am*PDH wild type and its variants were conducted. The results could rationalize experimental findings on an atomic resolution or suggest new experiments.

Chapter 3 describes a pure computational study employing MD simulations of wild type *Am*PDH. In the course of this study, the thermodynamic properties of D-glucose (GLC) binding were investigated by calculating conformational entropies and interaction energies within *Am*PDH. Additionally, important interactions between GLC and active site residues were computed. As a results of this study, several residues were selected for subsequent mutagenesis studies *in vitro* and *in silico*, which are described in detail in **Chapter 5**.

Chapter 4 describes a pure computational study to investigate the substrate promiscuity of wild type *Am*PDH employing MD simulations and free energy calculations. Modifications made to the GLC-topology allowed to sample all 32 possible aldohexopyranose sugars (SUG) within a single MD simulation. First, the solvation free energies of the 32 SUG in water could be computed. Second, the relative binding free energies of SUG to *Am*PDH were calculated, and good agreement with experiments was obtained. A combination of relative binding free energies and reweighted distance analysis rationalized experimentally detected SUG oxidation products and predicted new ones, which were not experimentally probed yet.

In **Chapter 5**, a combined experimental and computational approach was utilized to investigate structure–dynamics–function relationships of active site residues in *Am*PDH. First, active site variants as suggested in **Chapter 3** were generated and probed by biochemical and biophysical means. Results were compared to the wild type *Am*PDH and – where applicable – to MD simulations. The role of active site residues in GLC binding and turnover was investigated by steady-state kinetics and MD simulations. The GLC oxidation sites predicted by MD simulations were experimentally verified by GC-MS measurements. The preference of wild type *Am*PDH for C2 or C3 oxidized substrates as well as the effect of mutating His-556 to an alanine on the reductive

and oxidative half-reactions was tested *via* stopped-flow spectroscopy. Finally, the oxygen reactivities of all variants as well as the effect of the covalent FAD linkage on the redox potential were investigated.

Chapter 6 offers a conclusion and an outlook based on the studies that were performed in the course of this PhD thesis.

Abstract

The flavin-dependent sugar oxidoreductase pyranose dehydrogenase (PDH) from the plant litter-degrading fungus *Agaricus meleagris* oxidizes D-glucose (GLC) efficiently at positions C2 and C3. The closely related pyranose 2-oxidase (P2O) from *Trametes multicolor* oxidizes GLC only at position C2. Consequently, the electron output per molecule GLC is twofold for PDH compared to P2O making it a promising catalyst for bioelectrochemistry or for introducing novel carbonyl functionalities into sugars. The aim of this study was to rationalize the mechanism of GLC dioxidation employing molecular dynamics (MD) simulations of GLC-PDH interactions. Shape complementarity through nonpolar van der Waals interactions was identified as the main driving force for GLC binding. Together with a very diverse hydrogen-bonding pattern, this has the potential to explain the experimentally observed promiscuity of PDH towards different sugars. Based on geometrical analysis, we propose a similar reaction mechanism as in P2O involving a general base proton abstraction, stabilization of the transition state, an alkoxide intermediate, through interaction with a protonated catalytic histidine followed by a hydride transfer to the flavin N5 atom. Our data suggest the presence of the two potentially catalytic bases His-512 and His-556 increases the versatility of the enzyme, employing the most suitably oriented base depending on the substrate and its orientation in the active site. Our findings corroborate and rationalize the experimentally observed dioxidation of GLC by PDH and its promiscuity towards different sugars.

Authors

Michael Graf*, Urban Bren*, Dietmar Haltrich & Chris Oostenbrink (2013) Molecular dynamics simulations give insight into D-glucose dioxidation at C2 and C3 by *Agaricus meleagris* pyranose dehydrogenase. *J. Comput. Aided Mol. Des.* **27**, 295–304. *Both authors contributed equally to this work.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF): Doctoral Program BioToP—Biomolecular Technology of Proteins (FWF W1224), by the Vienna Science and Technology Fund (WWTF LS08-QM03), by the European Research Council (ERC, 260408), and the Slovenian Research Agency (ARRS; grant number P1-0002) which is gratefully acknowledged.

3.1 Introduction

The sugar oxidoreductase pyranose dehydrogenase (PDH, EC 1.1.99.29) is a monomeric flavoprotein of approximately 65 kDa containing ~7% carbohydrates. PDH is found in *Agaricaceae* and *Lycoperdaceae*, which represent a small group of fungi involved in lignocellulose breakdown from forest litter [1]. The enzyme was first isolated and characterized from *Agaricus bisporus* [2], and subsequently from *Macrolepiota rhacodes* [3], *Agaricus xanthoderma* [4] as well as *Agaricus meleagris* [1]. PDH from *A. meleagris* has been studied in most detail to date with respect to its biochemical properties and potential applications [5].

At commencement of this work, a preliminary version of the high resolution (1.6 Å) X-ray crystal structure of *A. meleagris* PDH with PDB code 4H7U was kindly made available to us [6]. Together with glucose oxidase (GOX, EC 1.1.3.4), the flavin domain of cellobiose dehydrogenase (CDH, EC 1.1.99.18), and pyranose 2-oxidase (P2O, EC 1.1.3.10), PDH belongs to the structural family of glucose–methanol–choline–oxidoreductases (GMC oxidoreductases) [5]. PDH exhibits broad carbohydrate substrate specificity compared to other GMC oxidoreductases. Depending on the sugar-substrate, it can perform (di)oxidations at C1, C2, C3, or C4 positions of the hexapyranose ring [5,7]. This property, however, depends strongly on the source of the enzyme as well as on the sugar substrate. For example, D-glucose (GLC) is (di)oxidized at positions C2 and C3 by PDH from *Agaricus* spp. but only at position C3 by PDH from *M. rhacodes*, while P2O oxidizes this sugar almost exclusively at C2 [8]. Consequently, the electron output per molecule of GLC is twofold for *Agaricus* PDH compared to P2O. This makes PDH a promising catalyst for applications in bioelectrochemistry or for introducing novel carbonyl functionalities into sugars [5].

In order to investigate the PDH-GLC interactions, molecular dynamics (MD) simulations were performed. MD simulations are becoming an increasingly popular standard tool in biosciences to explore dynamic system properties or to investigate features not readily accessible by experimental means. Because of this, MD simulations and experiments represent complementing methods to study different aspects of nature [9,10].

Unfortunately, no experimentally determined structure of any PDH-substrate complex is currently available. Hence the only possibility to study the detailed interactions of PDH with its major substrate GLC is by computational means. Although the GMC oxidoreductases PDH and P2O (PDB: 3LSK) [11] possess a relatively low overall sequence identity of 16%, their sugar-binding sites are very similar as demonstrated by superposition in **Figure 1A** with an atom-positional root-mean-square deviation (RMSD) of 0.13 nm for all heavy atoms in this figure. Moreover, experimentally determined P2O structures with GLC bound in two different orientations are available in the PDB: 3PL8 [12] and 2IGO [8]. Therefore, these two P2O-GLC structures were used to retrieve the coordinates of two PDH-GLC complexes for this study. After initial superposition, the GLC coordinates were manually grafted into the PDH structure in an orientation according to the P2O-PDB code 3PL8 (**Figure 1B**) [12] – termed pose A – or according to the P2O-PDB code 2IGO (**Figure 1C**) [8] – termed pose B.

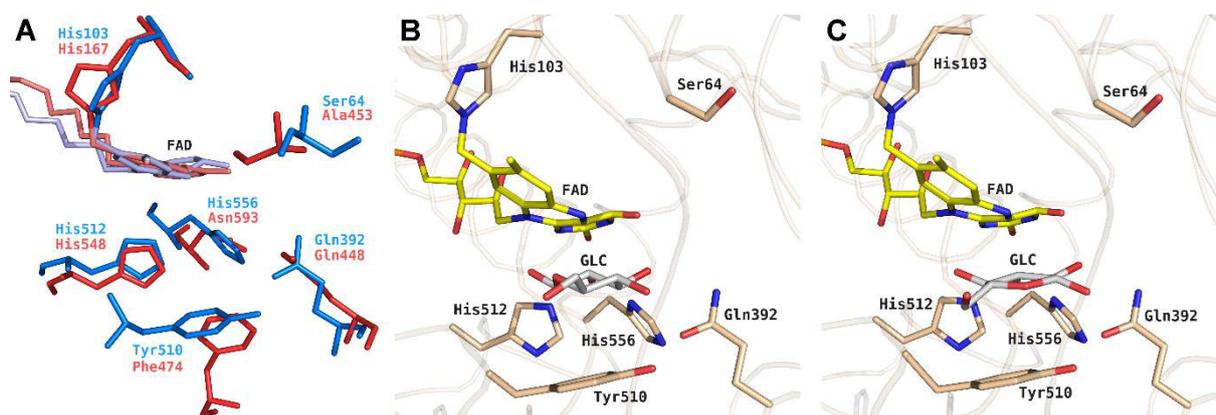


Figure 1: Active site structures of homotetrameric pyranose 2-oxidase (P2O) from *Trametes multicolor* and monomeric pyranose dehydrogenase (PDH) from *Agaricus meleagris*. (A) Superposition of active site residues from PDH, blue (PDB: 4H7U) and P2O, red (PDB: 3LSK). For clarity, the FAD moiety was colored light blue in PDH or light red in P2O and the atoms of the FAD ribityl side chains were omitted. (B, C) After superpositioning, D-glucose coordinates were grafted from the P2O structure with PDB accession codes (B) 3PL8 or (C) 2IGO into PDH. For clarity, P2O is not shown. Atom-coloring scheme: carbon (beige, protein; yellow, FAD; white, ligand), nitrogen (blue), oxygen (red), and phosphate (orange). Figures were generated using PyMOL (<http://www.pymol.org/>).

In contrast to P2O, PDH contains two, potentially catalytic residues, His-512 and His-556, in the active site (**Figure 1A**). In the present study, extensive MD simulations were conducted of PDH in its apo form, of GLC (D-glucose), and of a variety of complexes between the two, differing in the poses of the substrates and in the protonation state of the active site histidines. The results were compared in terms of important interactions, conformational entropies and interaction energies between PDH and GLC, and the structural dynamics and stability was followed in detail. An atomistic understanding of these properties will provide detailed insights into PDH-GLC interactions, which immediately suggest future site-directed mutagenesis experiments and can ultimately pave the way towards the desired bioelectrochemical applications.

3.2 Methods

3.2.1 Preparation of initial structures

A preliminary version of the crystal structure of *A. meleagris* PDH at 1.6-Å resolution with PDB entry code 4H7U without bound GLC served as a starting point [6]. In this structure, the isoalloxazine ring of the flavin is modified by a covalent monoatomic oxygen species at position C(4a). Since PDH does not react with oxygen under standard reaction conditions, this monoatomic oxygen species is most likely a result of oxygen radicals present during X-ray data collection and was therefore not considered for subsequent MD simulations. This structure does not contain the 25 amino acid long signal sequence (MLPRVTKLNSRLLSLALLGIQIARG), which is cleaved off

during secretion. Under physiological conditions, PDH is a glycosylated protein and sugar moieties are observed bound to Asn-75 and Asn-294, which were verified as potential glycosylation sites using the program NetNGlyc 1.0 [13]. In the current work, the sugar moieties NAG-901 (covalently attached to Asn-75), NAG-902 (covalently attached to Asn-294), and NAG-903 (covalently attached to NAG-902) were removed from the structure. Phosphate ion PO4-910, most likely a crystallization-buffer artifact surrounded by amino acid residues Arg-87 to Asp-90 and Pro-406 to Lys-407, was removed as well. The amino and carboxy termini were charged; all arginines, cysteines and lysines were protonated, and all aspartates and glutamates were deprotonated. Three different protonation states of His-512 and His-556 in the active site were considered: (i) His-512 being fully protonated and His-556 in its neutral state (proton at N ϵ), from now on labeled as PN; (ii) His-512 in its neutral state (proton at N δ) and His-556 protonated, from now on labeled as NP; (iii) both His-512 and His-556 fully protonated, from now on labeled as PP. The selection of the tautomeric states for neutral histidines was such that in the x-ray structure, the deprotonated nitrogen atoms pointed towards the active site. All remaining histidines were doubly protonated with exception of His-103, which is covalently bound to the FAD and was N δ -protonated. Two different substrate poses in the protein-substrate complex were generated: (i) Pose A: PDH alignment with P2O in complex with 3-fluoro-3-deoxy-D-glucose (PDB: 3PL8) [12]; (ii) Pose B: PDH alignment with P2O in complex with 2-fluoro-2-deoxy-D-glucose (PDB: 2IGO) [8] (**Figure 1B and C**). In all systems, the sugar coordinates were grafted into the active site of PDH after superposition and the fluorine of the sugar was replaced by a hydroxyl group. The combination of three protonation states and two substrate poses led to the definition of six protein-substrate complex systems. In addition, the apo protein was simulated (system PDH) using protonation state PP and system GLC was prepared consisting of β -D-glucose with coordinates taken from P2O-PDB 2IGO [8]. **Table 1** gives an overview of all simulated systems.

Table 1: Overview of simulated systems.

Code	Protonation (His-512 / His-556)	Ligand	Water molecules	Sodium ions	Total number of atoms	Runs
GLC ^a		D-glucose ^b	1,162	–	3,503	1
PN_A	+ / 0	D-glucose ^c	13,202	5	45,289	2
PN_B	+ / 0	D-glucose ^b	13,212	5	45,319	2
NP_A	0 / +	D-glucose ^c	13,198	5	45,277	2
NP_B	0 / +	D-glucose ^b	13,203	5	45,292	2
PP_A	+ / +	D-glucose ^c	13,770	4	45,244	2
PP_B	+ / +	D-glucose ^b	13,210	4	45,313	4
PDH	+ / +	–	13,434	4	45,968	1

^aD-glucose free in solution

^bD-glucose coordinates according to PDB code 2IGO

^cD-glucose coordinates according to PDB code 3PL8

3.2.2 Simulation setup

All MD simulations were carried out employing the GROMOS11 software package [14] with the 53A6 force field [15]. His-103 and FAD were covalently bound to each other and their force field parameters and topologies were adapted accordingly. The four studied systems were energy-minimized *in vacuo* using the steepest-descent algorithm: first, the sugar atoms were energy minimized with constrained PDH coordinates, and second, both the sugar- and the PDH atoms were energy minimized. Each energy-minimized system was placed into a periodic, pre-equilibrated, and rectangular box of SPC water [16]. Minimum solute to box-wall and minimum solute to solvent distances were set to 0.8 and 0.23 nm, respectively. To relax unfavorable atom-atom contacts between the solute and the solvent, energy minimization of the solvent was performed while keeping the solute positionally restrained using the steepest-descent algorithm. Finally, four to five water molecules that had the most favorable electrostatic potential for replacement by a positive ion were replaced by sodium ions to achieve electroneutrality in the protein systems.

For the equilibration, the following protocol was used: initial velocities were randomly assigned according to a Maxwell-Boltzmann distribution at 50 K. All solute atoms were positionally restrained through a harmonic potential with a force constant of $2.5 \times 10^4 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ not to disrupt the initial conformation, and the systems were propagated for 20 ps. In each of the five subsequent 20 ps MD simulations, the positional restraints were reduced by one order of magnitude and the temperature was increased by 50 K. Subsequently, the positional restraints were removed, rototranslational constraints were introduced on all solute atoms [17], and the systems were further equilibrated each for 20 ps at 300 K. In the end, a simulation at a constant pressure of 1 atm was conducted for 300 ps.

After equilibration, production runs of 10 ns at constant pressure (1 atm) and temperature (300 K) were carried out. Pressure and temperature were kept constant using the weak-coupling scheme [18] with coupling times of 0.5 and 0.1 ps, respectively. The isothermal compressibility was set to $4.575 \times 10^{-4} \text{ kJ}^{-1} \text{ mol nm}^3$, and two separate temperature baths were used for solute and solvent. The SHAKE algorithm was used to constrain bond lengths [19] allowing for 2-fs time-steps. Nonbonded interactions were calculated using a triple range scheme. Interactions within a short-range cutoff of 0.8 nm were calculated at every time step from a pairlist that was updated every fifth step. At these points, interactions between 0.8 and 1.4 nm were also calculated explicitly and kept constant between updates. A reaction field [20] contribution was added to the electrostatic interactions and forces to account for a homogenous medium outside the long-range cutoff using a relative dielectric constant of 61 as appropriate for the SPC water model [21]. Coordinate- and energy trajectories were stored every 0.5 ps for subsequent analysis.

3.2.3 Conformational entropy calculations

Conformational entropy calculations were performed according to the formulation of Schlitter [22]:

$$S_{\text{Schlitter}} = \frac{1}{2} k_B \ln \det \left[\mathbf{1} + \frac{k_B T e^2}{\hbar^2} \mathbf{M} \boldsymbol{\sigma} \right] \quad (1)$$

where k_B is Boltzmann's constant, T the absolute temperature, e Euler's number, \hbar Planck's constant divided by 2π , \mathbf{M} the $3N$ -dimensional diagonal matrix containing the N atomic masses of the solute atoms for which the entropy is calculated, and $\boldsymbol{\sigma}$ the covariance matrix of atom-positional fluctuations with the elements:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (2)$$

where x_i are the Cartesian coordinates of the atoms considered in the entropy calculation after a least-squares fit of the trajectory configurations using a particular subset of atoms.

3.3 Results and discussion

3.3.1 Structure of PDH

During all 10 ns MD simulations involving the protein, the root-mean-square deviation (RMSD) of the backbone atoms with respect to their initial crystal structure conformation remain below 0.3 nm, although for simulations NP_B, PP_A and PP_B the RMSD continues to rise, see

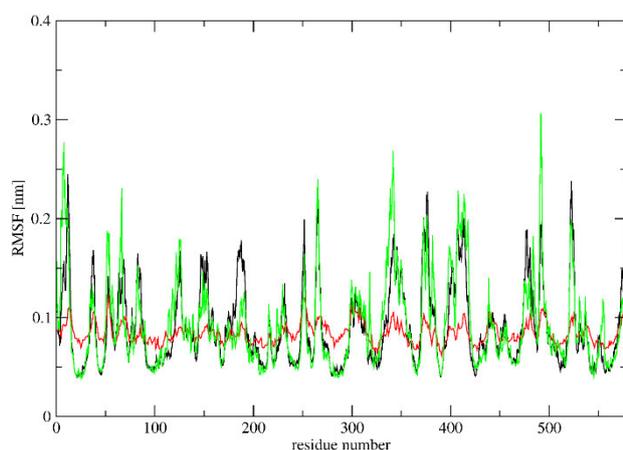


Figure 2: Root-mean-square fluctuations (RMSF) of the PDH backbone atoms. The RMSF values calculated from simulations PN_A, PN_B, NP_A, NP_B, PP_A, PP_B, were united according to the arithmetic mean (black curve) and compared to system PDH (green curve). RMSF values calculated from the crystal-structure B-factor values are depicted in red.

Supplementary Content Figure S1. Nevertheless, the observed RMSD values are low for all systems and indicate a stable protein backbone for all 10 ns MD simulations.

Secondary structure elements (DSSP) according to the Kabsch-Sander rules [23] for all simulations with PDH were assigned as a function of simulation time. Representative examples are compared in **Supplementary Content Figure S2**. Overall, the secondary structure elements are very well conserved during the simulations. The most prominent differences in DSSP between the different MD simulations can be observed for amino acid residues 500-512 comprising Tyr-510, Val-511, and His-512, which all directly interact with

GLC. The MD simulation of PDH shows a mixture between the β -bridge and the β -strand for these residues. In the complexes, the most prevalent secondary structure element for this region is either the β -bridge in e.g. system PP_A, simulation 2 or the β -strand in e.g. system PP_A, simulation 1. The manually inserted GLC has either slightly destabilizing or stabilizing effects on the PDH binding site and a clear trend cannot be identified.

Root-mean-square fluctuations (RMSF) of the PDH backbone were calculated over the 10 ns MD simulations (**Figure 2, Supplementary Content Figure S3**). As compared to the PDH run, all regions with high RMSF values for the complex simulations can be assigned to flexible solvent-exposed regions of the protein: Gln-145 to His-155, Asp-180 to Asn-190, Leu-345 to Asn-350, Lys-408 to Ala-415, and Met-473 to Lys-477. Note that the simulated RMSF are in good qualitative agreement with the RMSF calculated from the crystal-structure B-factor values (**Figure 2**).

3.3.2 Stability of the ligand, conformational entropies, and interaction energies

To investigate the thermodynamics of GLC binding, Schlitter ligand conformational entropies $S_{Schlitter}^L$ were calculated (**Table 2**). When fitted to GLC, $S_{Schlitter}^L$ for the complexes are lower compared to GLC simulated in water, indicating a loss of conformational entropy for GLC upon protein binding. The loss of conformational entropy upon binding is more pronounced for simulations PN_B and PP_A suggesting a more stable nature.

Table 2: Thermodynamics of binding between GLC and PDH in complex A and complex B. Schlitter $S_{Schlitter}^L$ ligand conformational entropies and average ligand-surrounding electrostatic $\langle V_{ES}^{L-S} \rangle$ as well as average ligand-surrounding van der Waals $\langle V_{vdW}^{L-S} \rangle$ interaction energies are reported as averages over individual simulation runs.

System	$S_{Schlitter}^{L,a,b}$ J mol ⁻¹ K ⁻¹	$S_{Schlitter}^{L,b,c}$	$\langle V_{ES}^{L-S} \rangle_d$ kJ mol ⁻¹	$\langle V_{vdW}^{L-S} \rangle_d$
GLC	–	318	-285 ± 1	-9.5 ± 0.2
PN_A	430 ± 8	241 ± 17	-265 ± 4	-47 ± 1
PN_B	344 ± 35	211 ± 25	-215 ± 3	-73 ± 1
NP_A	412 ± 15	243 ± 4	-254 ± 3	-53 ± 1
NP_B	413 ± 37	249 ± 21	-288 ± 6	-47 ± 1
PP_A	340 ± 5	213 ± 6	-267 ± 2	-58 ± 1
PP_B	417 ± 37	253 ± 27	-253 ± 3	-57 ± 1

^aCalculation performed after a rototranslational fit on the protein backbone atoms

^bError estimates calculated as standard deviations between individual simulations

^cCalculation performed after a rototranslational fit on D-glucose

^dError estimates calculated from block averaging [28] over individual runs and error propagation

In order to monitor the stability of GLC in the MD simulations of complex A and complex B, the root-mean-square deviation (RMSD) of the ligand was calculated with respect to its initial position after a translational and rotational fit of the protein backbone (**Supplementary Content Figure S4**). The RMSD for GLC involving protonation states NP and PP remains predominantly below 0.25 nm. Larger deviations are observed for the PN protonation state and to some extent simulations PP_B. To ensure that the increased mobility of GLC that was observed in these simulations was reproducible, two additional PP_B simulations were performed compared to the other systems, leading to very similar substrate mobility.

The values of $S_{Schlitter}^L$ that were obtained after a conformational fit on the protein backbone (second column in **Table 2**), do not only contain the conformational entropy, but also include the translational and rotational freedom of the substrate in the active site. Strikingly, these values confirm the stability of the substrate in systems PP_A, but also suggest a large structural stability for system PN_B. This suggests for the latter system that a stable pose was sampled, which was shifted by approximately 0.35 – 0.45 nm with respect to the initial structure.

The average electrostatic and van der Waals ligand-surrounding interaction energies of GLC in water and in the complex simulations are also listed in **Table 2**. The binding energies reveal that shape complementarity through nonpolar van der Waals interactions is the main driving force for GLC binding to PDH. This finding corroborates the experimentally observed promiscuity of the enzyme for a number of different sugar substrates [5]. In contrast, electrostatic contributions to binding between GLC and PDH are mostly unfavorable as indicated by higher average ligand-protein electrostatic interaction energies in complex A and complex B when compared to free GLC. However, electrostatics is still important for proper GLC orientation in the active site as indicated by the presence of different H-bonds in both complexes that will be discussed in detail in the subsequent section. A similar behavior was observed for the inhibition of extremely promiscuous cytochrome P450 enzymes [24,25]. Strikingly, the most favorable van der Waals interaction energies correlate with the least favorable electrostatic interaction energies for system PN_B, for which a shift in position was previously deduced. Further, the electrostatic interaction energies seem very comparable in size for the fully protonated systems (PP) and the singly protonated systems (PN, NP).

3.3.3 Important interactions between PDH and GLC

For the closely related GMC enzyme P2O, the reaction mechanism of the reductive half reaction involving a hydride transfer from atom C2 (GLC) to atom N5 (FAD) has been experimentally confirmed [26]. In order for such a hydride transfer to occur, the respective HC atom of GLC needs to be in the vicinity of the N5 atom of FAD. Therefore, the distances between atoms HC1-HC4 (GLC) and atom N5 (FAD) were monitored during the MD simulations of the complexes. Because GROMOS applies a united atom approach for methyldyne groups, virtual H-atoms were

introduced to calculate the distance between the respective H-atom and N5. **Figure 3** shows the distance distributions for the simulations. For complex A, the HC2-N5 distance is consistently the shortest and below a postulated cutoff of 0.3 nm in 15, 24 and 85% of the time in systems PN_A, NP_A and PP_A, respectively. The distance HC4-N5 is shorter than 0.3 nm for short periods of time as well, while the distance for neither HC1-N5 nor HC3-N5 is ever below 0.3 nm. We therefore conclude that GLC oriented according to pose A will be most likely oxidized at the C2 position. Put differently, pose A represents the C2 oxidation mode of GLC with respect to the N5 atom of the FAD. For pose B, the HC3-N5 distance distribution is the shortest and is below the chosen 0.3 nm cutoff in 88, 14 and 22% of the time in simulations PN_B, NP_B and PP_B, respectively. The HC1-N5 distance also drops below this cutoff occasionally, while none of the other hydrogen atoms are close enough to the N5 atom of FAD for a hydride transfer to occur. Therefore, we conclude that GLC oriented according to pose B represents the C3 oxidation mode of GLC.

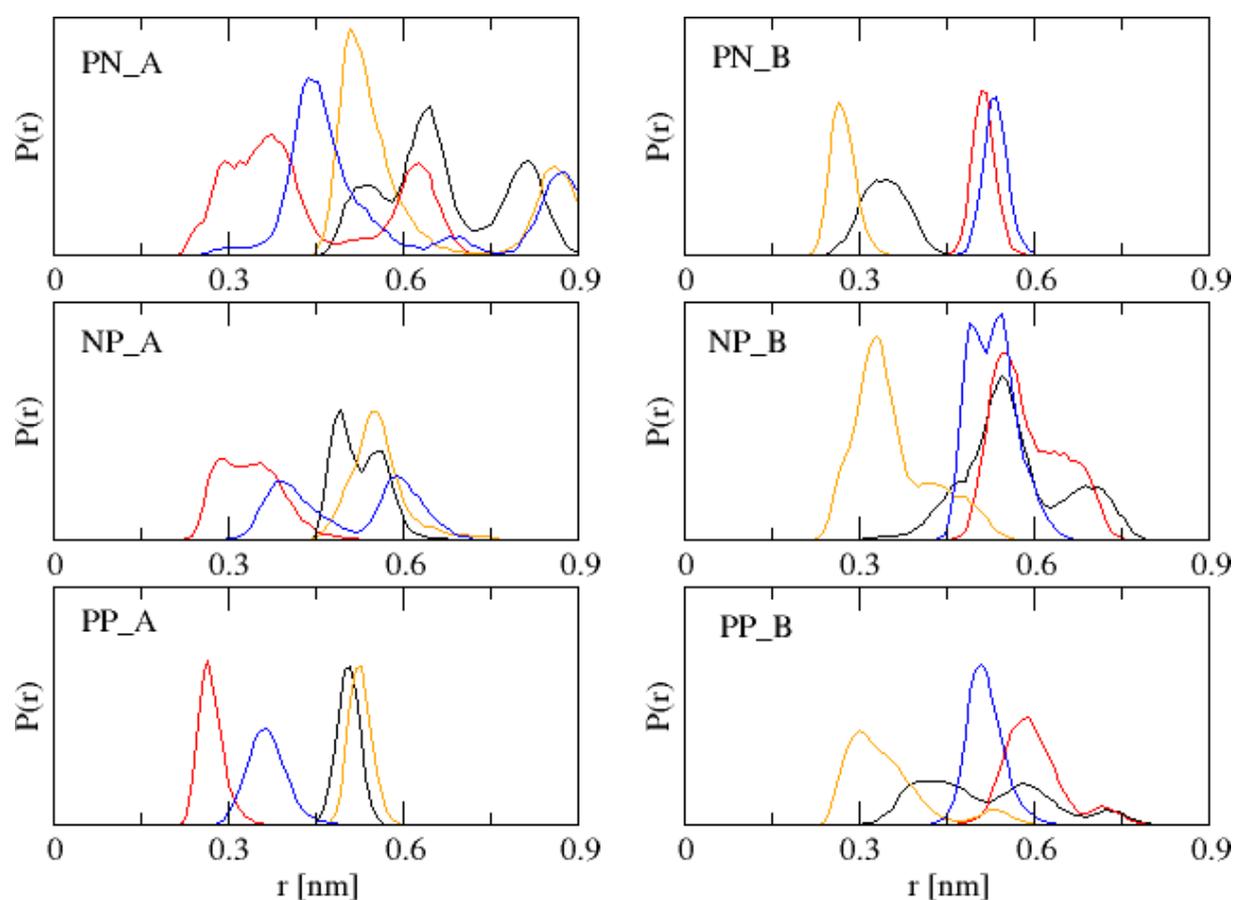


Figure 3: Normalized distance distributions between HC-atoms of D-glucose and the N5 atom of FAD over the combined PN_A, PN_B, NP_A, NP_B, PP_A, PP_B simulations. Coloring scheme: HC1-N5 (black), HC2-N5 (red), HC3-N5 (orange), HC4-N5 (blue).

The presence of hydrogen bonds was monitored utilizing a geometric criterion. A hydrogen bond was considered to be present if the hydrogen-acceptor distance was lower than 0.25 nm and the donor-hydrogen-acceptor angle was larger than 135°. Hydrogen bonds between PDH and GLC were monitored (**Table 3** and **Figure 4**).

Table 3: Occurrence of H-bonds between GLC and PDH, complex A, or complex B, averaged over all simulations. In brackets, the interacting atoms in GLC are indicated.

ID	Partner	PN_A (%)	PN_B (%)	NP_A (%)	NP_B (%)	PP_A (%)	PP_B (%)
1	Ser-64	12 (O6)					22 (O1)
2	Gly-105	40 (O3)					
3	Gly-359	13 (O4)					
4	Gln-392	21 (O2)	108 (O1/O5)	23 (O2)	24 (O2)	20 (O4)	30 (O1/O2)
5	Tyr-510	87 (O1/O2)		19 (O1)			33 (O3/O4)
6	Val-511		101 (O1/O2/O3)			181 (O1/O6)	29 (O3/O4)
7	His-512			32 (O1)	15 (O3)		
8	His-556		99 (O1/O2)		31 (O3)		
9	FAD	26 (O3)		31 (O2)	43 (O2/O3)		11 (O4)
	Total	199	308	105	113	201	125

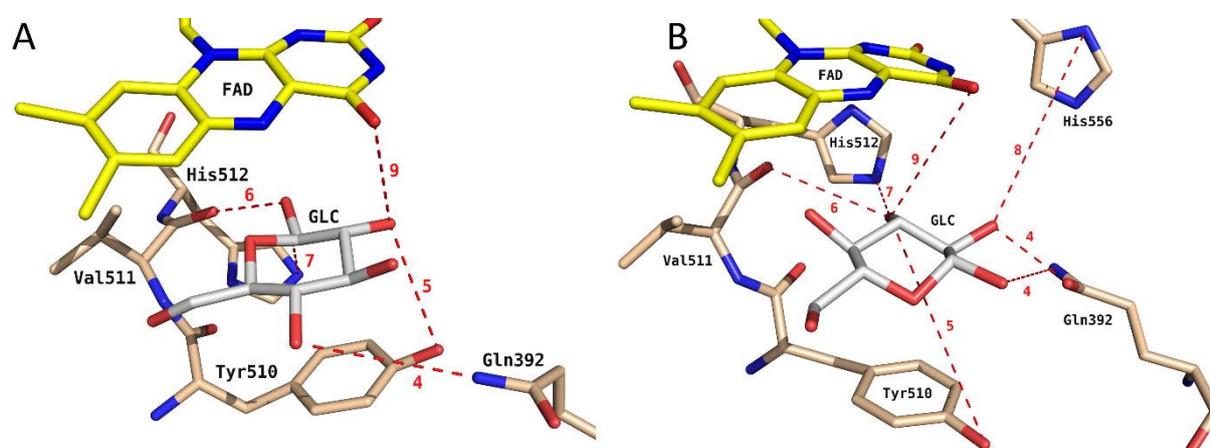


Figure 4: Two representative snapshots from MD trajectories of complex A (A) and complex B (B). The hydrogen bond network between D-glucose and amino acids involved in hydrogen bonding in either of the two complexes is shown. Hydrogen bonds are labeled according to their ID in **Table 3**. For clarity, hydrogen bonds with an occurrence below 15% are omitted. Coloring scheme as in **Figure 1**.

The largest amount of overall substrate-protein hydrogen bonds is observed for system PN_B, which was previously seen to have the largest loss of conformational entropy, the least favorable electrostatic interaction energy and the most favorable van der Waals interaction energy. It forms prominent hydrogen bonds with the sidechain of Gln-392, the backbone of Val-511 and the sidechain of His-556. Residues Gln-392 and Val-511 are involved in (partially very strong)

hydrogen bond interactions in the other systems as well. After system PN_B, systems PN_A and PP_A show the largest amount of hydrogen bonds. While this is due to a very strong interaction with the backbone of Val-511 for system PP_A, it is the result of a larger amount of shorter-lived hydrogen bonds for PN_A. This is again in agreement with the observed Schlitter entropies in **Table 2**; in system PP_A a single conformation is sampled, resulting in low entropy values, while system PN_A samples the active site more extensively, leading to higher entropy estimates. The versatile hydrogen-bonding pattern of systems PN_A, NP_A, NP_B and PP_B indicates once more the versatility of PDH in forming favorable interactions with different sugars.

A more detailed analysis of the observed hydrogen bonds offers an explanation for the experimental observation that the C2 product is not formed for the substrate methyl- α -D-glucopyranoside [27]. The methyl group at the O1 atom of the substrate removes the hydrogen bond donating capacity of this group. This reduces the total number of hydrogen bonds for systems PN_A, NP_A and PP_A by 63, 50 and 98%, respectively, and for systems PN_B, NP_B and PP_B by 46, 0, and 34%. Clearly, a loss of the hydrogen bond donating capacity of O1 in GLC destabilizes the C2 binding mode and leads only to C3 oxidation of methyl- α -D-glucopyranoside.

In the reaction mechanism elucidated for C2 oxidation by P2O, a general base initially abstracts the C2-OH proton of GLC. In a second step, the protonated His-548 stabilizes the appearing alkoxide intermediate at C2-O⁻ and acts as the catalytic residue [26]. In contrast, PDH exhibits two histidines in the active site (His-512 and His-556). Therefore, we calculated the distance distributions from the N δ or N ϵ atoms (whichever was closer) of His-512 and His-556 to atoms O2 or O3 of GLC in pose A and pose B, respectively (**Figure 5**). Again using a cutoff of 0.3 nm, it seems that both histidines can be positioned close enough to the hydroxyl groups to stabilize the deprotonation of the substrate. In systems PP_A and PP_B, the histidines both carry a positive charge and likely repel each other, leading to a more distinct preference for His-556 in pose A and for His-512 in pose B. Hydrogen bonds were observed between His-512 and His-556 in 35, 86, 78 and 68% of the time for systems PN_A, PN_B, NP_A and NP_B, respectively. Being connected through hydrogen bonds, the exact protonation state is likely to interconvert quite easily. Accordingly, for these systems, both histidine residues are able to interact with the GLC hydroxyl groups (**Figure 5**). Potentially, this offers another explanation for the substrate promiscuity of PDH. The ability of both histidines to take on the role of the catalytic base increases the versatility of the enzyme. In the system for which the GLC seems most stably anchored (PN_B), His-512 seems more likely to play the role of the catalytic residue. On the other hand, His-556 is involved in a stronger hydrogen bond interaction with the substrate. To a lesser extent, the opposite behavior is observed in system NP_A; although His-512 shows some hydrogen bonding with the substrate, His-556 comes slightly closer to atom O2, although a catalytic activity of His-512 cannot be excluded. Overall, we can conclude that His-512 samples distances that are in agreement with a catalytic function in all systems except for PP_A. His-556 is involved in more hydrogen bonds, but samples slightly less distances that are in agreement with a catalytic function. This observation is in agreement with the role of the two active site histidines His-502 and His-546 in aryl-alcohol oxidase

(AAO, PDB: 3FIM) [29], the structurally most similar GMC member to PDH [6]. In AAO, His-502 (corresponding to His-512 in PDH) plays a key role in the reductive half reaction acting as the catalytic base, whereas His-546 (corresponding to His-556 in PDH) has a more modest role, by correctly positioning the substrate through H-bonds during the reductive half reaction [29].

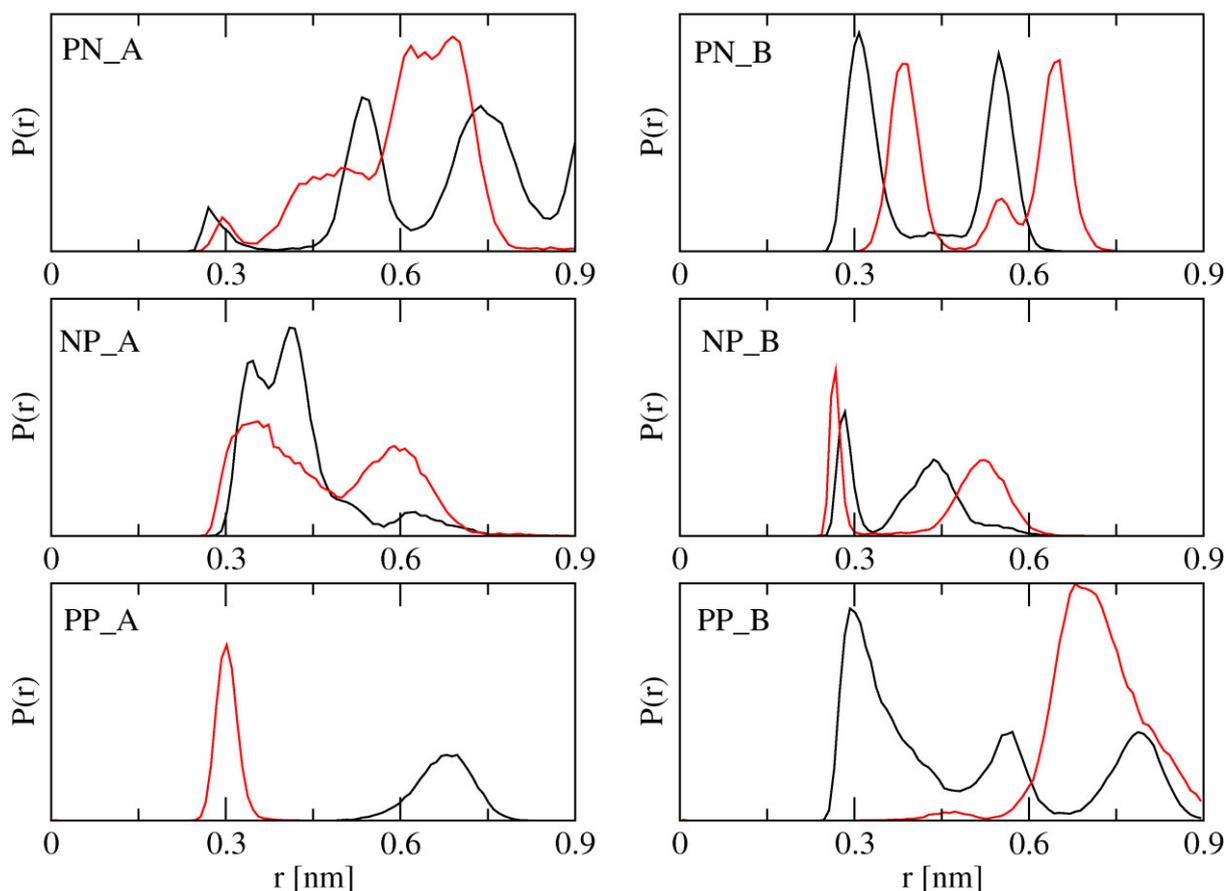


Figure 5: Normalized distributions of the shortest distance between N_δ or N_ϵ of His-512 (black) and His-556 (red) and D-glucose atom O2 in complex A (PN_A, NP_A, PP_A) and atom O3 in complex B (PN_B, NP_B, PP_B) in the combined simulations.

To conclude, **Figure 6** illustrates the proposed reaction mechanism of the first half reaction, corresponding to C2 and C3 oxidation of GLC. The two experimentally observed oxidation sites, corresponding to poses A and B in the simulations, can undergo deprotonation and subsequent stabilization involving both His-512 and His-556, as indicated in this figure.

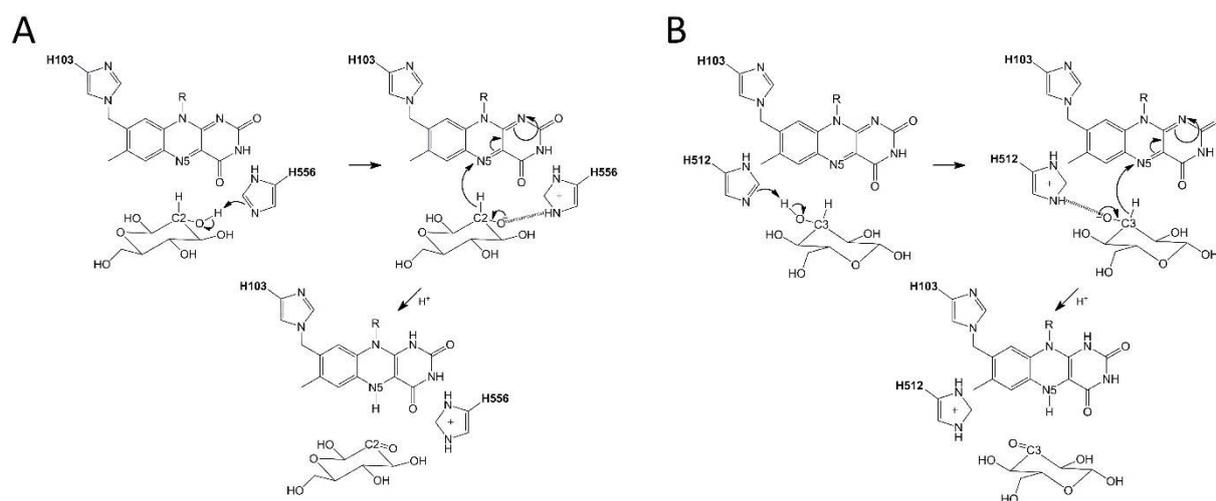


Figure 6: Proposed reaction mechanism for D-glucose oxidation at C2 in pose A (A) and at C3 in pose B (B). Both His-512 and His-556 can act as catalytic bases, with a slight preference for His-556 in the case of C2-oxidation (pose A) and for His-512 for C3-oxidation (pose B). The two suggested schemes are in agreement with the reaction mechanism of the reductive half-reaction of the structurally related P2O [26].

3.4 Conclusion

Extensive molecular dynamics simulations were performed to address the observed dioxidation of D-glucose at positions C2 and C3 by pyranose dehydrogenase as well as its promiscuous nature towards many different sugar substrates. RMSF analysis revealed good agreement between simulated and experimentally derived RMSF values. In addition, DSSP diagrams and RMSD values point toward a stable protein and GLC in all simulated systems providing additional confidence in the performed simulations.

To investigate the thermodynamics of GLC binding, the conformational entropies of GLC were calculated. Compared to GLC freely simulated in water, a loss of conformational entropy revealed an entropically unfavorable contribution to GLC binding for both complexes. Analysis of binding energies revealed that shape complementarity through nonpolar van der Waals interactions represents the main driving force for GLC binding, thereby corroborating the experimentally observed promiscuity of PDH. In contrast, electrostatic interactions were slightly unfavorable for GLC binding but they were still found to be important for proper GLC orientation within the active site through the hydrogen-bonding patterns that are observed.

A detailed hydrogen bond analysis offered an explanation for the absence of C2 product for the substrate methyl- α -D-glucopyranoside [27]: the methyl group at the O1 atom in this substrates prevents the donation of hydrogen bonds, which destabilizes the GLC C2 oxidation pose. Consequently, only pose B and the corresponding C3 oxidation product of GLC occurs.

Analysis of the distance distributions between GLC and active site histidines or the reactive N5 atom of FAD revealed insights into the proposed reaction mechanism. For pose A, the distances

between HC2 (GLC) and N5 (FAD) were smallest and in agreement with a possible hydride transfer reaction. Based on the distances between the N-atoms in histidines 512 and 556 and the O2 atom in GLC, a slight preference for His-556 as the catalytic residue was observed, although His-512 could easily take over this function. For complex B, the distance between HC3 (GLC) and N5 (FAD) were smallest and a slight preference for His-512 as the catalytic residue was observed. These findings suggest that complex A represents the C2-oxidation mode, while complex B represents the C3-oxidation mode. The versatility of the enzyme is possibly enhanced by the presence of two histidine residues in the active site, which can both take on the role of the catalytic residue. To conclude, our data suggests a similar reaction mechanism previously reported for P2O [26]: oxidation of either C2 or C3 of GLC is accomplished through a proton abstraction by a general base and transition state stabilization by the active site His-556 or His-512, followed by a hydride transfer to atom N5 of FAD.

In summary, the promiscuity of PDH with respect to other GMC oxidoreductases can be attributed to various effects. Not only does PDH catalyze both C2 and C3 oxidation of GLC, while P2O only catalyzes C2 oxidation, PDH is also characterized by oxidation of a wide range of different carbohydrates. First, the active site offers various hydrogen bonding possibilities, such that alternative substrates or poses can be expected to find favorable interactions as well. Second, the actual binding is governed mostly by non-specific van der Waals interactions, allowing for the binding of many different substrates. Third, the presence of two active site histidines expands the versatility from the binding to the catalytic process itself. The P2O enzyme does not only have a substrate loop, increasing the selectivity through specific interactions, it also lacks the second histidine residue observed in PDH.

As a final note, our work suggests various experiments that will be performed in the near future to characterize the function of several active site residues. Obviously, site directed mutagenesis on His-512 and His-556 may confirm our hypothesis that both residues may have a catalytic role. Further, based on the observed hydrogen bonding patterns, we suggest mutating Gln-392, Tyr-510 and Val-511, as these may influence the preferred product formation. Even though the interactions with Val-511 are through its backbone, a mutation to a more bulky, hydrophobic residue should be able to disrupt the observed hydrogen bonds. These findings indicate that MD simulations are indeed developing towards a standard tool from which experimentalists can draw ideas for new investigations.

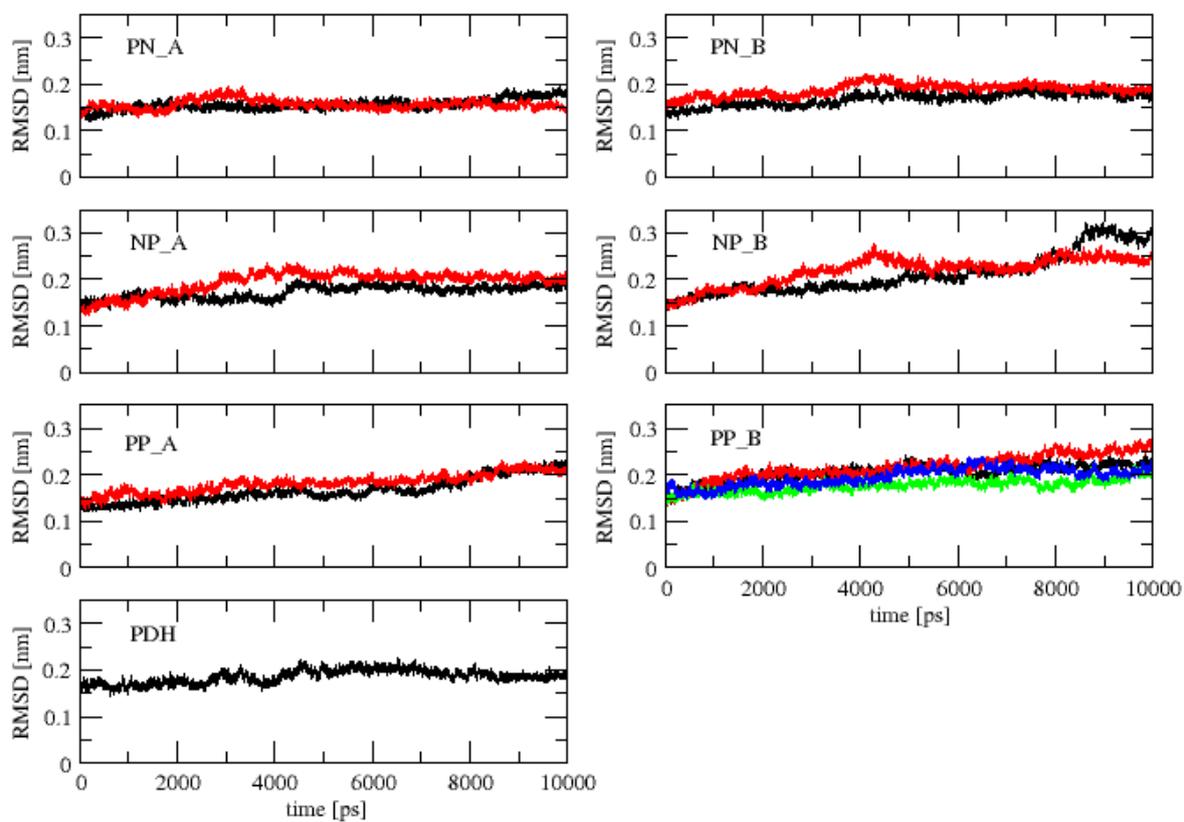
3.5 References

- 1 Sygmund C, Kittl R, Volc J, Halada P, Kubátová E, Haltrich D & Peterbauer CK (2008) Characterization of pyranose dehydrogenase from *Agaricus meleagris* and its application in the C-2 specific conversion of D-galactose. *J. Biotechnol.* **133**, 334–342.
- 2 Volc J, Kubátová E, Wood DA & Daniel G (1997) Pyranose 2-dehydrogenase, a novel sugar oxidoreductase from the basidiomycete fungus *Agaricus bisporus*. *Arch. Microbiol.* **167**, 119–125.
- 3 Volc J, Kubátová E, Daniel G, Sedmera P & Haltrich D (2001) Screening of basidiomycete fungi for the quinone-dependent sugar C-2/C-3 oxidoreductase, pyranose dehydrogenase, and properties of the enzyme from *Macrolepiota rhacodes*. *Arch. Microbiol.* **176**, 178–186.
- 4 Kujawa M, Volc J, Halada P, Sedmera P, Divne C, Sygmund C, Leitner C, Peterbauer C & Haltrich D (2007) Properties of pyranose dehydrogenase purified from the litter-degrading fungus *Agaricus xanthoderma*. *FEBS J.* **274**, 879–894.
- 5 Peterbauer CK & Volc J (2010) Pyranose dehydrogenases: biochemical features and perspectives of technological applications. *Appl. Microbiol. Biotechnol.* **85**, 837–848.
- 6 Tan TC, Spadiut O, Wongnate T, Sucharitakul J, Krondorfer I, Sygmund C, Haltrich D, Chaiyen P, Peterbauer CK & Divne C (2013) The 1.6 Å crystal structure of pyranose dehydrogenase from *Agaricus meleagris* rationalizes substrate specificity and reveals a flavin intermediate. *PLoS ONE* **8**, e53567.
- 7 Leitner C, Volc J & Haltrich D (2001) Purification and characterization of pyranose oxidase from the white rot fungus *Trametes multicolor*. *Appl. Environ. Microbiol.* **67**, 3636–3644.
- 8 Kujawa M, Ebner H, Leitner C, Hallberg BM, Prongjit M, Sucharitakul J, Ludwig R, Rudsander U, Peterbauer C, Chaiyen P, Haltrich D & Divne C (2006) Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J. Biol. Chem.* **281**, 35104–35115.
- 9 Karplus M & Kuriyan J (2005) Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6679–6685.
- 10 Van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, Gee P, Geerke DP, Glättli A, Hünenberger PH, Kastenzholz MA, Oostenbrink C, Schenk M, Trzesniak D, van der Vegt NFA & Yu HB (2006) Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Ed. Engl.* **45**, 4064–4092.
- 11 Tan T-C, Pitsawong W, Wongnate T, Spadiut O, Haltrich D, Chaiyen P & Divne C (2010) H-bonding and positive charge at the N(5)/O(4) locus are critical for covalent flavin attachment in *Trametes* pyranose 2-oxidase. *J. Mol. Biol.* **402**, 578–594.
- 12 Tan T-C, Haltrich D & Divne C (2011) Regioselective control of β -D-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy. *J. Mol. Biol.* **409**, 588–600.
- 13 Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S & Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633–1649.

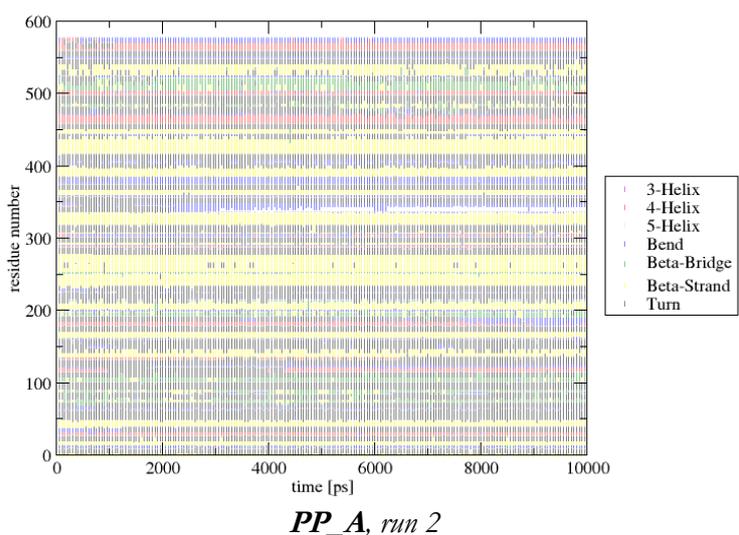
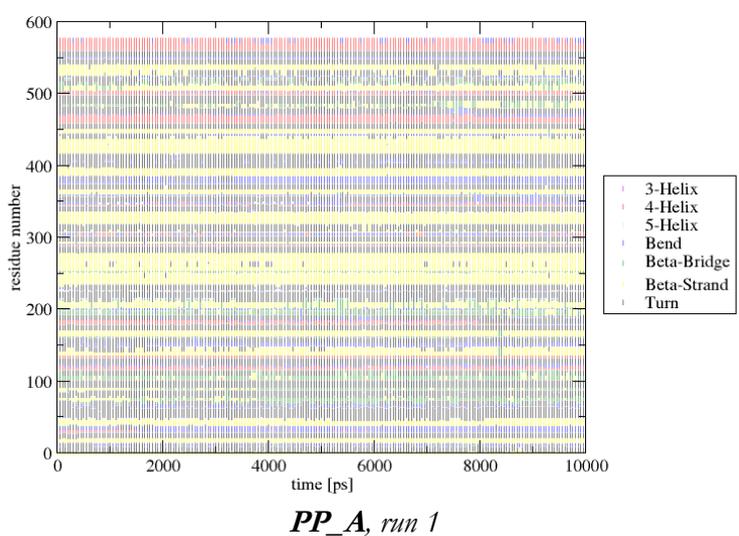
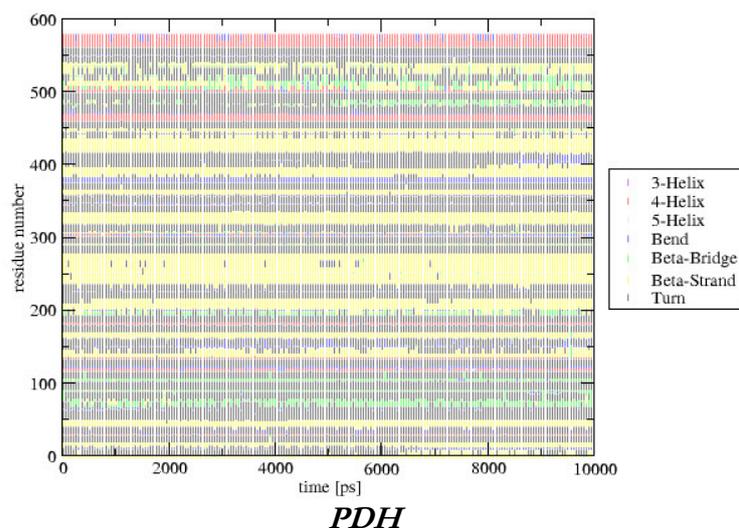
- 14 Schmid N, Christ CD, Christen M, Eichenberger AP & van Gunsteren WF (2012) Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comp. Phys. Commun.* **183**, 890–903.
- 15 Oostenbrink C, Villa A, Mark AE & Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676.
- 16 Berendsen HJC, Postma JPM, Van Gunsteren WF & Hermans J (1981) Interaction models for water in relation to protein hydration. *Intermolecular forces*. Reidel: Dordrecht, The Netherlands, pp. 331-342.
- 17 Amadei A, Chillemi G, Ceruso MA, Grottesi A & Di Nola A (2000) Molecular dynamics simulations with constrained roto-translational motions: Theoretical basis and statistical mechanical consistency. *J. Chem. Phys.* **112**, 9–23.
- 18 Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
- 19 Ryckaert J-P, Ciccotti G & Berendsen HJ. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* **23**, 327–341.
- 20 Tironi IG, Sperb R, Smith PE & van Gunsteren WF (1995) A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459.
- 21 Heinz TN, van Gunsteren WF & Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **115**, 1125–1136.
- 22 Schlitter J (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem. Phys. Lett.* **215**, 617–621.
- 23 Kabsch W & Sander C (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
- 24 Vasanthanathan P, Olsen L, Jørgensen FS, Vermeulen NPE & Oostenbrink C (2010) Computational prediction of binding affinity for CYP1A2-ligand complexes using empirical free energy calculations. *Drug Metab. Dispos.* **38**, 1347–1354.
- 25 Bren U & Oostenbrink C (2012) Cytochrome P450 3A4 inhibition by ketoconazole: Tackling the problem of ligand cooperativity using molecular dynamics simulations and free-energy calculations. *J. Chem. Inf. Model.* **52**, 1573–1582.
- 26 Wongnate T, Sucharitakul J & Chaiyen P (2011) Identification of a catalytic base for sugar oxidation in the pyranose 2-oxidase reaction. *ChemBioChem* **12**, 2577–2586.
- 27 Volc J, Sedmera P, Halada P, Daniel G, Příkladová V & Haltrich D (2002) C-3 oxidation of non-reducing sugars by a fungal pyranose dehydrogenase: Spectral characterization. *Mol. Cat. B*, **17**, 91–100.
- 28 Allen MP & Tildesley DJ (1989) Computer simulation of liquids. Clarendon Press: Oxford, England.

- 29 Hernández-Ortega A, Lucas F, Ferreira P, Medina M, Guallar V & Martínez AT (2012) Role of active site histidines in the two half-reactions of the aryl-alcohol oxidase catalytic cycle. *Biochemistry* **51**, 6595–6608.

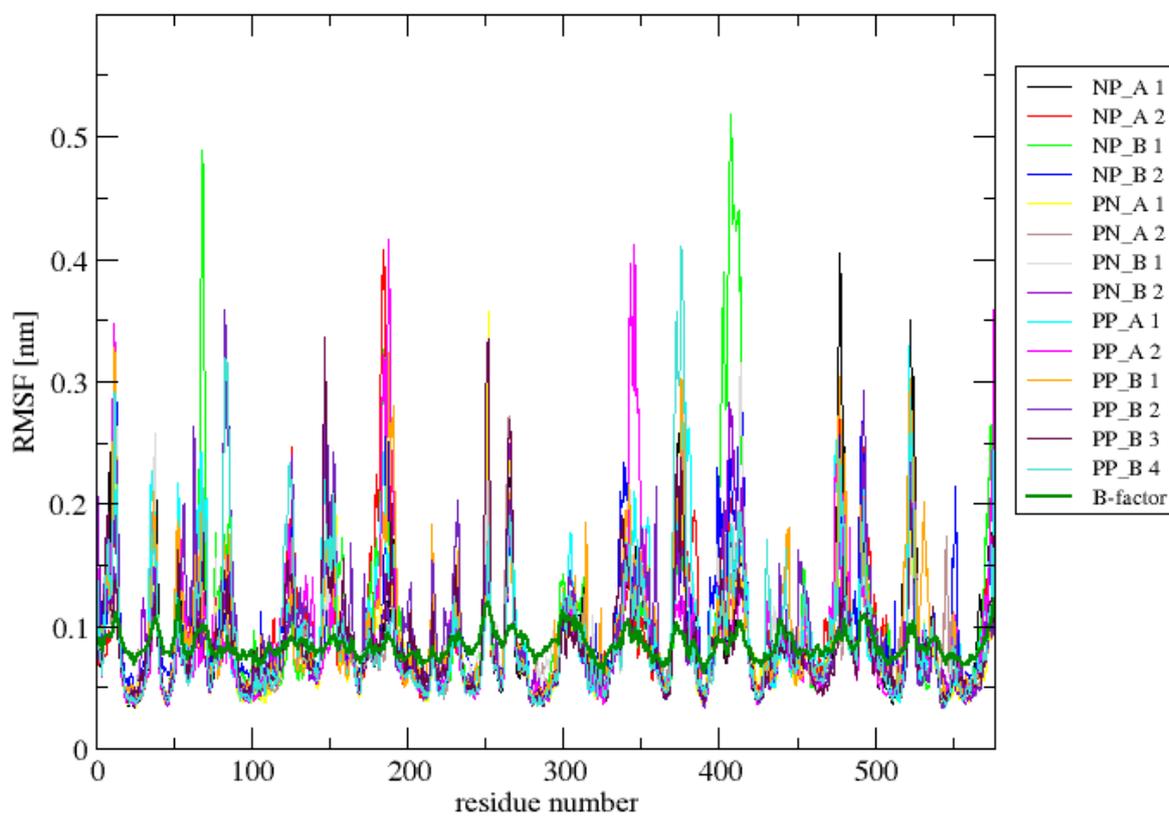
3.6 Supplementary material



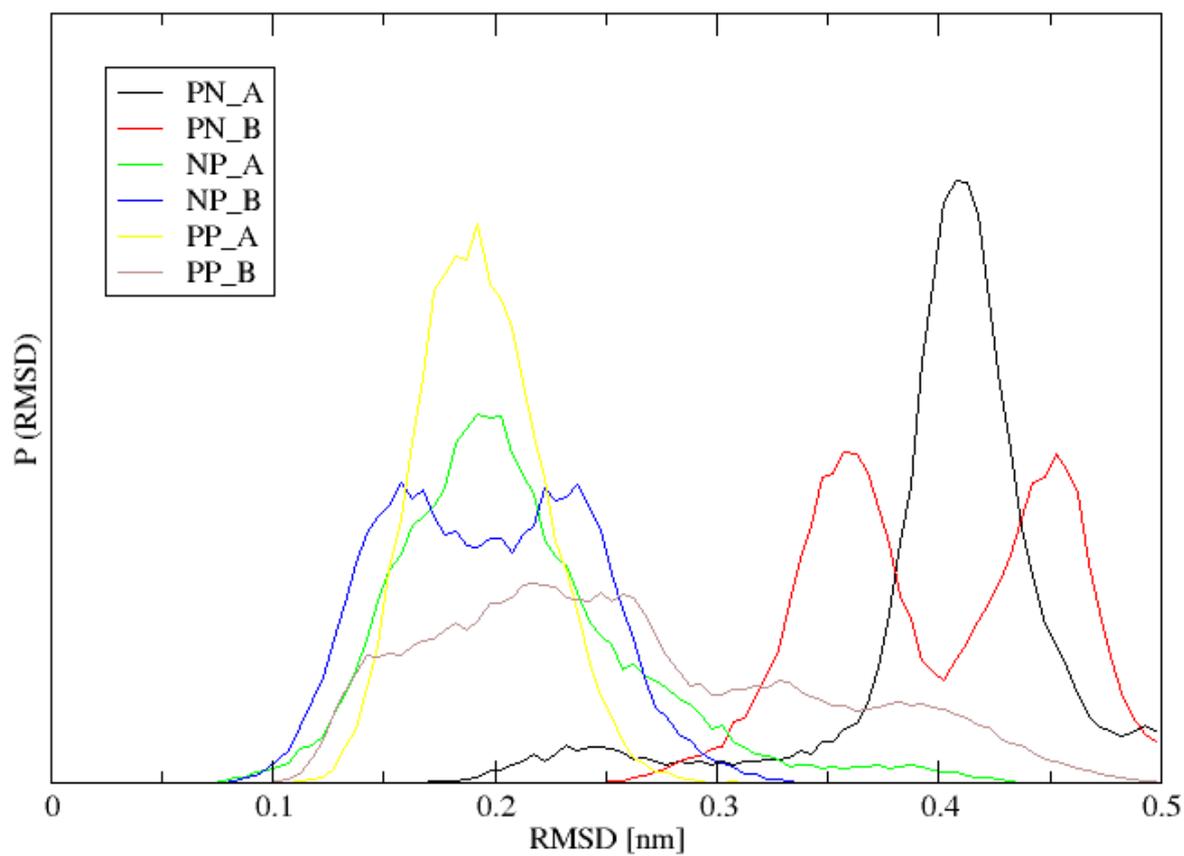
Supplementary Figure S1: Root-mean-square deviation (RMSD) of the backbone during the 10 ns MD simulations of systems PN_A, PN_B, NP_A, NP_B, PP_A, PP_B and PDH. Coloring scheme: run 1 (black), run 2 (red), run 3 (green), run 4 (blue).



Supplementary Figure S2: Secondary structure elements (DSSP) classification as a function of time for selected simulations.



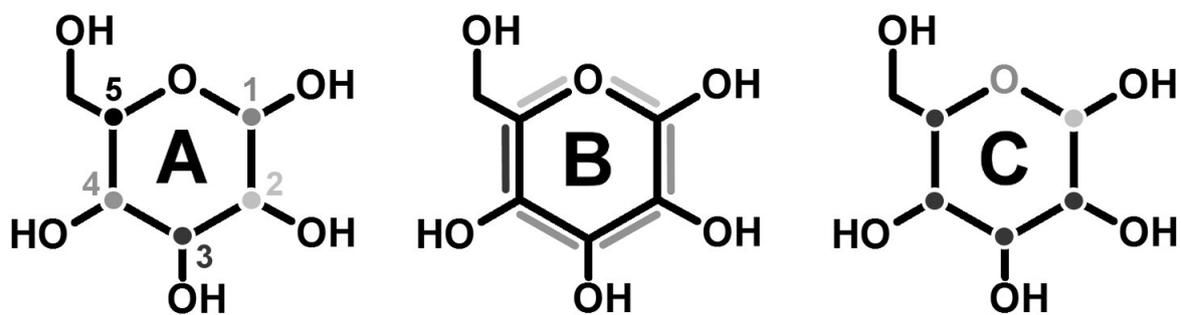
Supplementary Figure S3: *Atom-positional root-mean-square fluctuations (RMSF) for the individual simulations. The average of the simulation curves yields the curve in figure 2 of the main manuscript. The B-factors are indicated with a thicker green curve.*



Supplementary Figure S4: *Distributions of the atom-positional root-mean-square deviation (RMSD) of D-glucose after a translational fit of the protein backbone with respect to the initial structure during the 10 ns MD simulations.*

CHAPTER 4

Ligand Promiscuity of PDH



Abstract

The flavoenzyme pyranose dehydrogenase (PDH) from the litter decomposing fungus *Agaricus meleagris* oxidizes many different carbohydrates occurring during lignin degradation. This promiscuous substrate specificity makes PDH a promising catalyst for bioelectrochemical applications. A generalized approach to simulate all 32 possible aldohexopyranoses in the course of one or a few molecular dynamics (MD) simulations is reported. Free energy calculations according to the one-step perturbation (OSP) method revealed the solvation free energies (ΔG_{solv}) of all 32 aldohexopyranoses in water, which have not yet been reported in the literature. The free energy difference between β - and α -anomers ($\Delta G_{\beta-\alpha}$) of all D-stereoisomers in water were compared to experimental values with a good agreement. Moreover, the free-energy differences (ΔG) of the 32 stereoisomers bound to PDH in two different poses were calculated from MD simulations. The relative binding free energies ($\Delta\Delta G_{\text{bind}}$) were calculated and, where available, compared to experimental values, approximated from K_m values. The agreement was very good for one of the poses, in which the sugars are positioned in the active site for oxidation at C1 or C2. Distance analysis between hydrogens of the monosaccharide and the reactive N5-atom of the flavin adenine dinucleotide (FAD) revealed that oxidation is possible at HC1 or HC2 for pose A, and at HC3 or HC4 for pose B. Experimentally detected oxidation products could be rationalized for the majority of monosaccharides by combining $\Delta\Delta G_{\text{bind}}$ and a reweighted distance analysis. Furthermore, several oxidation products were predicted for sugars that have not yet been tested experimentally, directing further analyses. This study rationalizes the relationship between binding free energies and substrate promiscuity in PDH, providing novel insights for its applicability in bioelectrochemistry. The results suggest that a similar approach could be applied to study promiscuity of other enzymes.

Authors

Michael Graf, Lin Zhixiong, Urban Bren, Dietmar Haltrich, Wilfred van Gunsteren & Chris Oostenbrink (2014) Pyranose dehydrogenase ligand promiscuity: A generalized approach to simulate monosaccharide solvation, binding, and product formation. *PLoS Comput Biol* 10(12): e1003995.

Acknowledgments

The authors thank Prof. Alan Mark for fruitful discussions and useful suggestions, especially concerning Figure 4. This work was supported by the Austrian Science Fund (FWF W1224): Doctoral Program BioToP – Biomolecular Technology of Proteins, by the Swiss National Science Foundation (SNF 200020-137827), by the European Research Council (ERC grant numbers 228076 and 260408), by the Vienna Science and Technology Fund (WWTF LS08-QM03), and the Slovenian Research Agency (ARRS grant numbers P1-0002 and J1-5448) which is gratefully acknowledged.

4.1 Introduction

Generally, enzymes are perceived as being specific for both their substrates and the reactions they catalyze [1]. Deviations from such behavior are often seen as unwanted side effects or even errors in the biological function of the enzyme that come at an additional energetic cost for the organism. Although this feature has long been recognized to be useful in other contexts, for example in the recognition of multiple antigens by the same germline antibody [2–4], such enzymes are often characterized by poor overall catalytic efficiencies and termed promiscuous. Starting in 1976, this paradigm started to shift when Jensen drew a link between promiscuity and protein evolution [5]. He hypothesized that the first enzymes had very broad substrate specificities that evolved to more specialized forms via duplication, mutation, and selection of the corresponding genes. This was corroborated by later studies that investigated the evolutionary implications of promiscuity such as the adaptation of enzymes towards novel xenobiotics, e.g. halogenated compounds or antibiotics, in the course of a few decades [6,7]. Although systematic screens for promiscuous enzyme functions are not feasible because of the vast number of possible different substrates and reactions, there are many indications and examples that promiscuity is rather the rule than the exception [6]. Especially in the past two decades, enzyme promiscuity received considerable attention, and enzymes that can take over the function of related enzymes in an organism *via* their promiscuous activities have been extensively investigated [8–10]. These studies suggest that metabolic pathways are intertwined in many unexpected ways, which ultimately gives the organism a higher survival potential under changing environmental conditions. Regulation of such metabolic pathways as well as promiscuity itself at the gene-, transcript-, protein-, and localization-level and the associated reaction conditions are other thriving research areas [1,11]. Moreover, promiscuity is often observed for close homologs in protein families and distant homologs within superfamilies [11]. Individual family members have frequently evolved from a common ancestor through gene duplication and subsequent specialization. These members share the same fold and catalytic strategy, and consequently the main activity of one family member is often found as the promiscuous activity of another family member. Nobeli and coworkers refer to this phenomenon as ‘family’ promiscuity as opposed to ‘individual’ or ‘pure’ promiscuity, which is associated to multiple functions of a single enzyme [11]. The molecular mechanisms underlying promiscuity are manifold, including post-translational modifications, multiple domains, oligomeric states, protein flexibility, partial recognition, multiple interaction sites or a single site with diverse interacting residues, allosteric interactions, flexibility as well as size and complexity of the interaction partner, chemical scaffolds, and protonation states of active site residues [6,11,12]. Hydrophobic interactions, diverse hydrogen bonding, flexibility, and nonpolar van der Waals interactions combined with negligible electrostatics were found to be the main driving forces for promiscuity [11,13–15]. Consequently, understanding the molecular mechanisms and energetics leading to enzyme promiscuity is a valuable asset in the field of protein design and engineering as well as drug development, and therefore they have been investigated extensively [1,16]. In view of various causes and effects involving promiscuity, it is not surprising that the definition of the term is not exact and

combinations of different definitions occur [1,6,11,16,17]. In this article, the term ‘promiscuity’ is used in the context of relaxed substrate specificity [1,18] in order to perform similar chemical reactions on related substrates [17].

A prototypical example of ‘family’ promiscuity [11] can be found in the structural family of GMC oxidoreductases, named after three representatives utilizing either glucose, methanol, or choline as their substrate [19]. Although the four initially characterized GMC family members share only between 23 and 32 % sequence similarity and possess diverse catalytic activities with a wide range of substrate specificity, they share the same overall architecture and catalyze a similar chemical reaction, the oxidation of an alcohol moiety. Cavener speculates about an ancestral protein of this family that could bind to many different substrates [19], which it converted with low catalytic efficiencies while mutations paved the way for high specificities towards the individual substrates. A more recent addition to the GMC oxidoreductase family is the glycosylated enzyme pyranose dehydrogenase (PDH, EC 1.1.99.29), reacting with many different carbohydrates. It contains a monocovalently bound flavin adenine dinucleotide (FAD) and has a mass of approximately 65 kDa. Although PDHs from other sources with similar biochemical properties have been studied [20–22], the enzyme from *Agaricus meleagris* has been characterized most extensively so far: a wealth of biochemical data [13,23–25] and a high-resolution X-ray structure with PDB code 4H7U [26] are readily available. To date, its exact physiological role is unknown. However, because of the natural habitat of *Agaricus meleagris* on lignocellulose-rich forest litter and PDH’s reactivity with a multitude of different carbohydrates found during lignin degradation, PDH is most likely involved in lignocellulose breakdown [25]. Compared to other GMC oxidoreductases, it oxidizes many different aldoses and ketoses in pyranose form as well as heteroglycosides, glucooligosaccharides, sucrose, and lactose, which can be (di)oxidized at C1–C4. A comprehensive list of its impressively broad substrate spectrum can be found in the paper of Sedmera and coworkers [27] and an updated version in the review of Peterbauer and Volc [25]. The reactivity towards many different carbohydrate-substrates makes PDH a very interesting enzyme to study in the context of substrate-promiscuity. In this manuscript, we investigate computationally the promiscuous nature of PDH towards the pyranose form of monosaccharides that are turned over by the enzyme. All 32 possible combinations of α - and β -anomers as well as D- and L-stereoisomers of glucose, mannose, galactose, talose, allose, altrose, gulose, and idose will be considered.

Molecular dynamics (MD) simulations of PDH were applied to study the interactions with the monosaccharides described above. The aim of this study was to gain a deeper understanding of the promiscuous nature of PDH towards monosaccharides. This involved a generalized approach of extensive MD simulations and free energy calculations using the one-step perturbation (OSP) method [28,29] to calculate monosaccharide binding and solvation. The OSP method is an efficient means to obtain free-energy differences of similar molecules from a simulation of a carefully designed reference molecule for which the sampling is such that configurations are sampled that are representative of the molecules or states in which one is interested. In the past, OSP has successfully been applied to reproduce and predict binding free energies of a series of compounds

to a common receptor [29,30], to study the stereoselective binding of substrates to a promiscuous enzyme [31,32], to study conformational preferences of molecules that show slow transitions in regular simulations [33,34] or the effect of changes in force-field parameters on conformational equilibria [35,36].

In the current work, we investigate (i) the solvation free energies of the 32 above-mentioned monosaccharides in water; (ii) the free energy differences of the β/α -anomers of the D-stereoisomers in water; (iii) the relative binding free energies for all monosaccharides and (iv) the occurrence of reactive poses for all monosaccharides. Where experimental data were available, comparisons were made and good to excellent agreements were observed. Furthermore, our work offers predictions of properties that have not yet been described experimentally.

4.2 Methods

4.2.1 Preparation of initial structures

The structure preparations were essentially performed as reported previously [13]. In short, a preliminary version of the 1.6 Å resolution X-ray structure of PDH (PDB code 4H7U) served as starting point [26]. The covalent monoatomic oxygen species at C(4a), which is most likely an X-ray artifact, was removed. As a glycoprotein, the structure of PDH comprised covalently attached sugar moieties at surface residues Asn-75 and Asn-294. The influence of these glycosylated residues on the active site is expected to be negligible and consequently the glycan structures were removed. A PO_4^{3-} ion at the surface, which is most likely a crystallization buffer artifact, was removed as well. The amino and carboxy termini were charged; all arginines, cysteines and lysines were protonated, and all aspartates and glutamates were deprotonated. In our previous study, we propose that PDH oxidizes its sugar substrate *via* a general base proton abstraction [13], which requires one of the two active site histidines (His-512 and His-556) being neutrally charged. The most stable protonation state fulfilling this requirement was obtained when His-512 was fully protonated and His-556 was in its neutral state (proton at N_ϵ). The selection of the tautomeric state for the neutral His-556 was such that in the X-ray structure its deprotonated nitrogen atom pointed towards the active site. The remaining histidines were doubly protonated, except for His-103, which is covalently attached to the FAD and was protonated at N_δ . No structure of PDH comprising a monosaccharide-substrate in the active site was available at commencement of this work. Therefore, PDH and the closely related GMC oxidoreductase pyranose 2-oxidase (P2O, EC 1.1.3.10) were superimposed with an atom-positional root-mean-square deviation (RMSD) of 0.13 nm for all heavy atoms of their sugar-binding sites. Two different P2O structures were used, in which the bound sugar roughly differs in a 180° rotation around the axis going through C2 and C5 of the tetrahydropyran ring to allow for (di)oxidations at all possible sites (C1–C4). Superposition of PDH and P2O in complex with 3-fluoro-3-deoxy- β -D-glucose (PDB: 3PL8) [37] yielded pose A (**Figure 1A**),

whereas pose B (**Figure 1B**) was obtained by aligning PDH with P2O in complex with 2-fluoro-2-deoxy- β -D-glucose (PDB: 2IGO) [38].

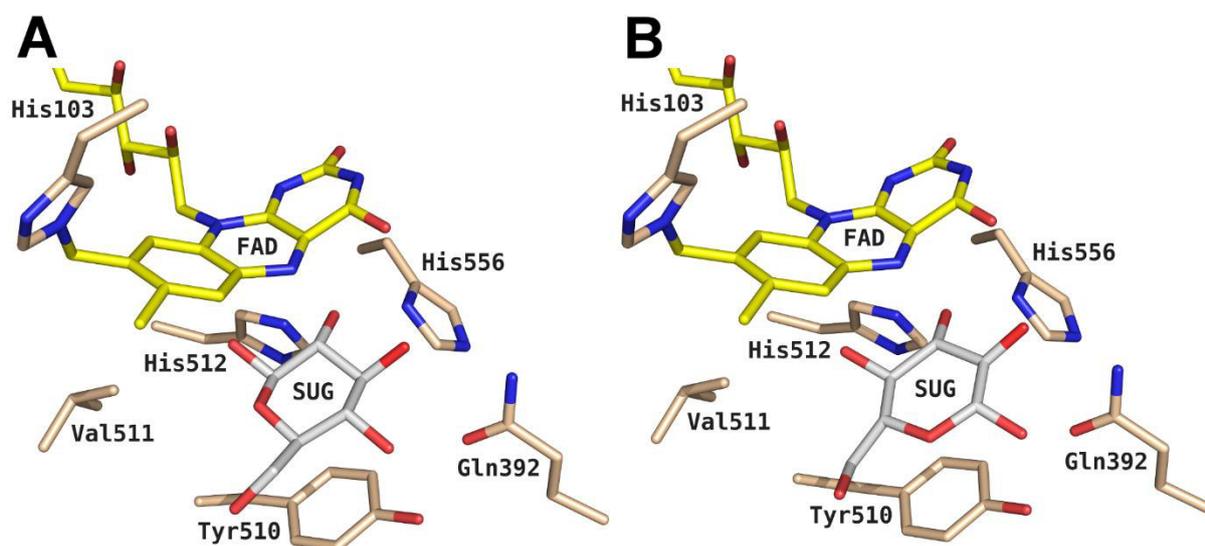


Figure 1: Active site of pyranose dehydrogenase (PDH; PDB code 4H7U) from *Agaricus meleagris* with bound sugar (SUG; here β -D-glucose). To transfer the sugar coordinates from pyranose 2-oxidase (P2O) from *Trametes multicolor* into PDH, the X-ray structures of both enzymes were first superimposed. The β -D-glucose coordinates from the P2O structures with PDB codes (A) 3PL8 – termed pose A – or (B) 2IGO – termed pose B – were grafted into PDH. Atom-coloring scheme: carbon (beige, protein; yellow, FAD; white, ligand), nitrogen (blue), and oxygen (red). The figure was generated using PyMOL (<http://www.pymol.org/>).

After grafting the monosaccharide coordinates into PDH's active site, the fluorine of the sugar was replaced by a hydroxyl group. This procedure ultimately yielded system PDH-SUG, with the monosaccharide bound to PDH according to pose A or pose B. For simulations of sugar without PDH, the coordinates of β -D-glucose from P2O-PDB 2IGO [38] were used. For the description of the interactions with the sugar, a united atom force field was used. Chirality around CH-groups in such a force field is enforced through an improper dihedral potential energy term. In order to allow transitions between equatorial and axial positions of the attached hydroxyl groups and to sample all 32 possible monosaccharides in a single MD simulation, following changes were made to the topology of β -D-glucose, following suggestions in references [28] and [31] as indicated in **Figure 2**:

- (a) Improper dihedral angle interactions at C1–C5 were turned off (**Figure 2A**) – leading to model SUG^a .
- (b) In addition to (a), the proper dihedral force constants (k_ϕ) for the ring torsional dihedral angles were lowered according to the coloring scheme in **Figure 2B** – leading to model SUG^{ab} .
- (c) In addition to (a) and (b), the bond angle bending force constants (k_θ) for the bond angles surrounding all ring atoms (C1–C5 and O) were lowered according to the coloring scheme in **Figure 2C** – leading to model SUG^{abc} .

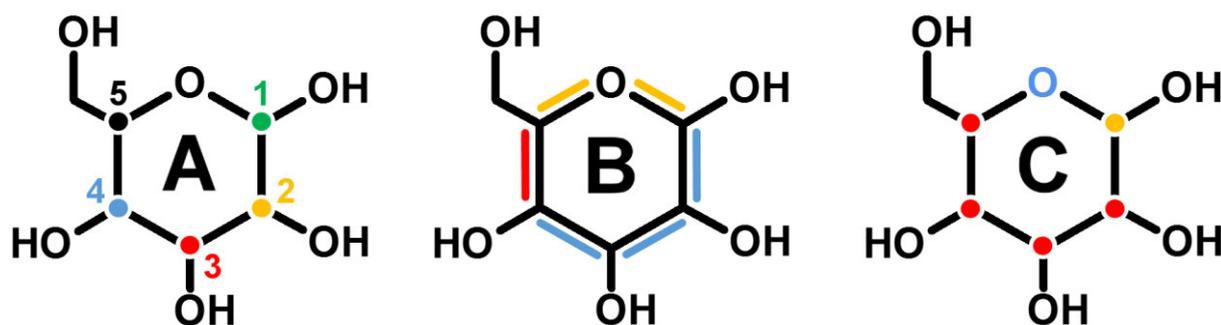


Figure 2: Modifications made to the topology of system SUG (also refer to **Supplementary Data**). (A) System SUG^a : improper dihedrals (ID) at stereocenters C1–C5 were turned off. Numbers indicate the name of the C-atom and the ID-position within the 5-digit ID code in **Table 1** and **Table 3**. Colors are in agreement with the coloring schemes in **Figures 5-7**: green (C1, ID1), yellow (C2, ID2), red (C3, ID3), blue (C4, ID4), and black (C5, ID5). (B) System SUG^{ab} : in addition to system SUG^a , proper dihedral force constants (k_ϕ) for the ring torsional dihedral angles were lowered according to the following coloring scheme: blue (two force constants from 2.09 to 0.418 kJ mol^{-1} and one from 5.92 to 1.05 kJ mol^{-1}), red (one force constant from 5.92 to 1.05 kJ mol^{-1}), and yellow (two force constants from 3.77 to 1.05 kJ mol^{-1}). (C) System SUG^{abc} : in addition to system SUG^{ab} , bond angle bending force constants (k_θ) for the bond angles surrounding the ring atoms (C1–C5 and O) were lowered according to following coloring scheme: blue (one bond angle from 380 to 285 kJ mol^{-1}), red (two bond angles from 320 to 285 kJ mol^{-1}), yellow (three bond angles from 320 to 285 kJ mol^{-1}).

The building block and the changes made to the β -D-glucose topology to define the three reference molecules are further detailed in the supplementary material (**Supplementary Table S1**).

4.2.2 Simulation setup

MD simulations were conducted using the GROMOS 11 software package [39] employing the 53A6 force field [40]. In this parameter set, carbohydrates are described according to the parameter set of [41], and the topology of β -D-glucose is given in the supplementary material. Note that the GROMOS force field is a united atom force field, which is crucial for the modifications of the sugar interactions described in the previous section. His-103 and FAD were covalently attached to each other and their topologies and force field parameters were adapted accordingly. All systems

were energy-minimized in vacuo employing the steepest-descent algorithm: for the PDH-SUG complexes, the sugar atoms were energy-minimized with constrained PDH coordinates after which both SUG and PDH atoms were energy-minimized. A 1- μ s stochastic dynamics (SD) simulation of SUG in vacuum was performed, referred to as SUG_{vac}, for which the energy-minimized structure of SUG was used. For MD simulations of SUG and the complex PDH-SUG in water (SUG_{water} and PDH-SUG, respectively) the structures were placed into a rectangular, periodic, and pre-equilibrated box of SPC water [42]. All water molecules within 0.23 nm of any solute atom were removed from the box and a minimum solute to box-wall distance of 0.8 nm was enforced. In order to relax unfavorable atom-atom contacts between the solute and the solvent, energy-minimization of the solvent was performed while keeping the solute positionally restrained using the steepest-descent algorithm. Finally, five water molecules with the most favorable electrostatic potential for replacement by a positive ion were substituted with sodium ions to achieve electroneutrality in systems PDH-SUG.

The following protocol was used to thermalize and equilibrate the system: initial velocities were randomly assigned according to a Maxwell-Boltzmann distribution at 50 K. All solute atoms were positionally restrained through a harmonic potential energy term with a force constant of 2.5×10^4 kJ mol⁻¹ nm⁻² in order not to disrupt the initial conformation, and the systems were propagated for 20 ps. In five subsequent 20 ps MD simulations, the positional restraints were reduced by one order of magnitude and the temperature was increased by 50 K. Subsequently, the positional restraints were removed, roto-translational constraints introduced on all solute atoms [43], and the systems were further equilibrated, each for 20 ps at 300 K. Finally, an equilibration at a constant pressure of 1 atm was conducted for 300 ps.

After equilibration, production runs at constant pressure (1 atm) and temperature (300 K) were performed. For the SUG_{water} systems, one production run of 100 ns was performed. For the PDH-SUG systems, each with SUG^a, SUG^{ab}, or SUG^{abc} bound according to pose A or pose B, two independent 50 ns production runs (termed md1 and md2) were conducted, leading to a total of 12 independent PDH-SUG simulations. Pressure and temperature were kept constant using the weak-coupling scheme [44] with coupling times of 0.5 and 0.1 ps, respectively. The isothermal compressibility was set to 4.575×10^{-4} kJ⁻¹ mol nm³, and two separate temperature baths were used for solute and solvent. The SHAKE algorithm was applied to constrain all solute bond lengths [45] as well as the solvent geometry in simulation SUG_{water}. Because of simulation efficiency, the SETTLE algorithm was applied to constrain solvent geometry [46] in system PDH-SUG. In all cases, constraining the bond lengths allowed for 2-fs time-steps. Nonbonded interactions were calculated according to a triple range scheme. Interactions within a short-range cutoff of 0.8 nm were calculated at every time-step from a pair list that was updated every fifth step. At these points, interactions between 0.8 and 1.4 nm were also calculated explicitly and kept constant between updates. A reaction field [47] contribution was added to the electrostatic interactions and forces to account for a homogenous medium outside the long-range cutoff using a relative dielectric constant

of 61 as appropriate for the SPC water model [48]. Coordinate and energy trajectories were stored every 0.5 ps for subsequent analysis.

4.2.3 Free energy calculations

The one-step perturbation (OSP) method relies on the application of Zwanzig's perturbation formula which is exact in the limit of infinite sampling [49]. In practice, the free-energy difference between a possibly unphysical reference molecule represented by Hamiltonian H_R and a physically relevant compound represented by Hamiltonian H_A is accurately estimated if a simulation of the reference molecule samples a sufficiently high number of configurations relevant for the real compound. In those cases the free energy can efficiently be calculated using

$$\Delta G_{AR} = -k_B T \ln \langle e^{-(H_A - H_R)/k_B T} \rangle_R \quad (1)$$

where the angular brackets $\langle \rangle_R$ indicate an ensemble average computed from the simulation of the reference state and $k_B T$ represents Boltzmann's constant multiplied by the absolute temperature. Since only the energy difference $H_A - H_R$ appears in the exponential, only the few energy terms that are different between the compounds need to be re-evaluated over the real compounds, here involving only the covalent interactions indicated in **Figure 2**.

The free energy differences ΔG_{AR} can subsequently be used to estimate various physically relevant free energy differences, such as the solvation free energies, relative to the reference state

$$\Delta \Delta G_{\text{solv,AR}} = \Delta G_{AR}(\text{water}) - \Delta G_{AR}(\text{vacuum}) \quad (2)$$

The free energy difference between α - and β -anomers of specific sugars can be computed as

$$\Delta \Delta G_{\beta-\alpha}(\text{water}) = \Delta G_{\alpha R}(\text{water}) - \Delta G_{\beta R}(\text{water}) \quad (3)$$

where the subscripts α and β refer to the α - and β -anomers of a single monosaccharide. The binding free energy relative to the reference compound is calculated as

$$\Delta \Delta G_{\text{bind,AR}} = \Delta G_{AR}(\text{PDH-SUG}) - \Delta G_{AR}(\text{water}) \quad (4)$$

and relative to another compound as

$$\Delta \Delta G_{\text{bind,AB}} = \Delta \Delta G_{\text{bind,AR}} - \Delta \Delta G_{\text{bind,BR}} \quad (5)$$

One limitation of the OSP approach is the fact that most simulation effort is spent on unphysical reference molecules, reducing the direct insight into the structure and dynamics of the real compounds. However, the ensemble average of any property Q for the real compounds may be obtained using [50,51]

$$\langle Q \rangle_A = \frac{\langle Q e^{-(H_A - H_R)/k_B T} \rangle_R}{\langle e^{-(H_A - H_R)/k_B T} \rangle_R} \quad (6)$$

which was used here to analyze the average occurrence of reactive poses for the real compounds. The distances between H-atoms HC1, HC2, HC3, and HC4 and the N5 atom of the FAD cofactor were calculated for the reference state simulations as r_{HN} . Consistent with our previous study [13], a particular conformation was considered as reactive for a specific carbon if the corresponding value of r_{HN} was below 0.3 nm, such that the average occurrence can be calculated as

$$\langle N(< 0.3 \text{ nm}) \rangle_{\text{R}} = \langle H(0.3 \text{ nm} - r_{\text{HN}}) \rangle_{\text{R}} \quad (7)$$

where $H(x)$ is the Heaviside step function, *i.e.* $H(x) = 1$ for $x \geq 0$ and $H(x) = 0$ for $x < 0$. By replacing Q in equation (6) by $H(0.3 \text{ nm} - r_{\text{HN}})$, we obtain the average amount of catalytically active poses of the real compounds.

In the current work, multiple reference compounds R were applied (SUG^{a} , SUG^{ab} , and SUG^{abc}) whereas individual estimates were combined by transferring the free energy estimates to a common reference state. One can easily show that expressing the ensemble average for reference compound $R1$ of equation (1) as an umbrella-weighted ensemble, calculated from a simulation of reference state $R2$ using equation (6), can be expressed as

$$\Delta G_{\text{AR1}} = \Delta G_{\text{AR2}} - \Delta G_{\text{R1R2}} \quad (8)$$

where both terms on the right-hand side are readily calculated from the simulation of $R2$. This way, simulations of the three reference states lead to three estimates of ΔG_{AR1}^i , which can be exponentially averaged to obtain

$$\Delta G_{\text{AR}} = -k_{\text{B}}T \ln \overline{e^{-\Delta G_{\text{AR}}^i/k_{\text{B}}T}} \quad (9)$$

where the overbar indicates an average over three values of i [52]. Statistical error estimates for the individual ensemble averages used in equation (1) were obtained from covariances and the statistical inefficiency as described in [53]. The uncertainty in a series of N correlated sequential observations x_n , with expectation value $\langle x \rangle = 1/N \sum_n x_n$, becomes $\delta^2 = [\langle x_n^2 \rangle - \langle x_n \rangle^2]/(N/g)$ where g is the statistical inefficiency, defined as $g = 1 + 2\tau$, with τ the auto-correlation time of the normalized autocorrelation function, $C_t = [\langle x_n x_{n+t} \rangle - \langle x_n \rangle^2]/[\langle x_n^2 \rangle - \langle x_n \rangle^2]$ [53]. The individual error estimates of ΔG_{AR}^i were subsequently weighted by $e^{-\Delta G_{\text{AR}}^i/k_{\text{B}}T}$ to obtain the statistical uncertainty on ΔG_{AR} .

4.3 Results and discussion

4.3.1 Monosaccharide solvation

To find a suitable reference state, which is crucial for reliable free energy calculations according to the OSP method, MD simulations of system SUG_{water} with changes to the topology according to SUG^a , SUG^{ab} , and SUG^{abc} were conducted. As a typical example, **Figure 3** shows the distributions of the improper dihedral angle 5 (ID5) centered on atom C5 for the three 100 ns MD simulations. For SUG_{water}^a (black), ID5 is not evenly distributed and samples mostly the region around +30 degrees. SUG_{water}^{ab} (red) and SUG_{water}^{abc} (blue) both show more equal distributions, indicating that both stereo-configurations are equally sampled. To use a topology with minimal changes with respect to the real compounds, SUG^{ab} was selected as the most suitable reference state in water. Similarly, SUG^{ab} was used for the 1 μs SD simulation in vacuo (SUG_{vac}^{ab}). In contrast, SUG^a , SUG^{ab} , and SUG^{abc} (**Figure 2**) were all selected for simulations and analysis of system PDH-SUG, in order to sample as many stereoisomers as possible. Consequently, 12×50 ns MD simulations of system PDH-SUG were conducted: three different SUG topologies, two different SUG binding poses (pose A and pose B) and two independent simulations for each (md1 and md2). MD simulations of system PDH-SUG will be referred to as *e.g.* PDH-SUG $_{\text{md1}}^{abc}$ (pose A).

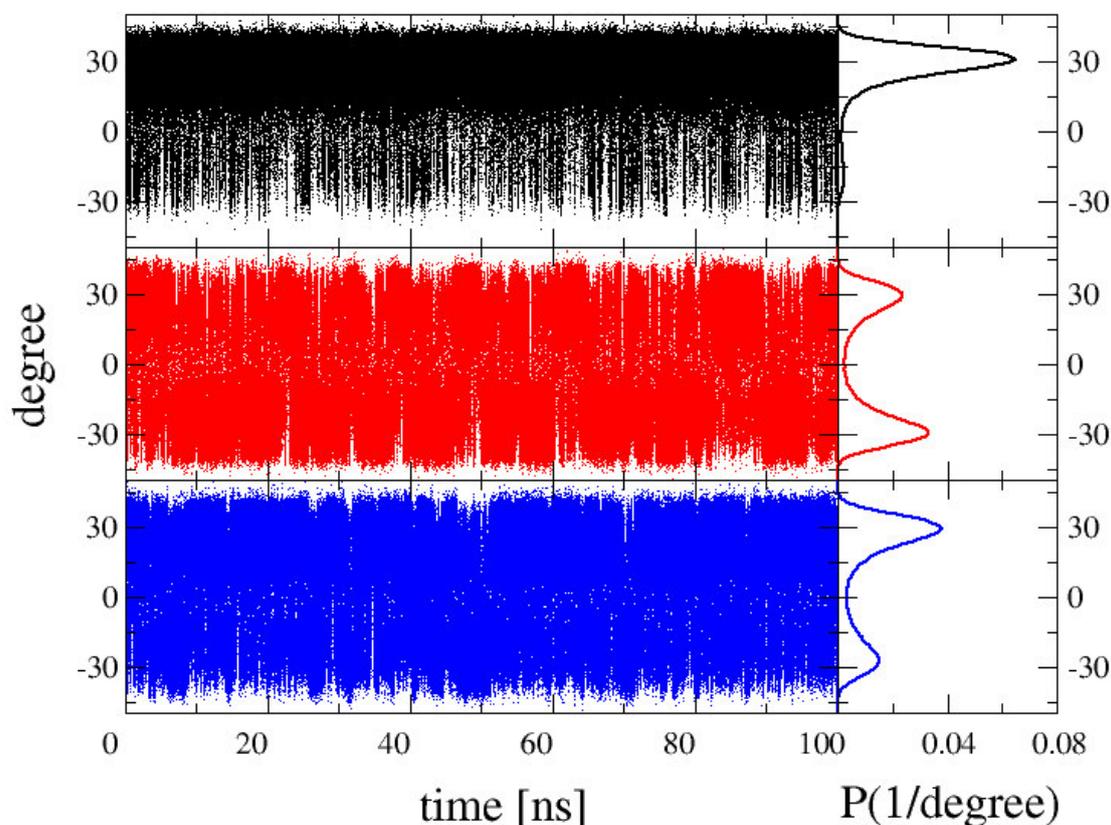


Figure 3: Distributions of the improper dihedral 5 for the three 100 ns MD simulations of system SUG_{water} with changes to the topology according to **Figure 2**. Coloring scheme: SUG^a (black), SUG^{ab} (red), and SUG^{abc} (blue).

Table 1: Relative free-energy differences (ΔG) of the SD or MD simulations for the 32 sugars with respect to the reference state in vacuum (SUG_{vac}) or water (SUG_{water}) and their calculated relative free energies of solvation ($\Delta\Delta G_{solv}$) in kJ/mol. The digits ‘2’ or ‘4’ of the improper dihedral code in this table corresponds to the improper dihedral type code listed in the IMPDIHEDRAL-block of the SUG-topology (**Supplementary Table S1**). The sequence of the digits ‘2’ and ‘4’ follows the improper dihedral angles of the sugar according to **Figure 2A**.

Sugar code	Improper dihedral code	Sugar name	ΔG [kJ/mol]		$\Delta\Delta G_{solv}$ [kJ/mol]
			SUG_{vac}^{ab}	SUG_{water}^{ab}	
1	22222	β -D-glucose	26.7 ± 0.2	31.7 ± 0.9	5.0 ± 0.9
2	24222	β -D-mannose	32.4 ± 0.3	37.9 ± 1.0	5.6 ± 1.0
3	22242	β -D-galactose	33.1 ± 0.3	34.5 ± 1.3	1.5 ± 1.3
4	24242	β -D-talose	31.8 ± 0.3	40.8 ± 2.5	9.1 ± 2.5
5	42222	α -D-glucose	20.5 ± 0.1	34.4 ± 1.0	13.9 ± 1.0
6	44222	α -D-mannose	29.3 ± 0.1	33.9 ± 0.5	4.5 ± 0.5
7	42242	α -D-galactose	25.3 ± 0.1	36.2 ± 1.2	10.9 ± 1.2
8	44242	α -D-talose	30.1 ± 0.1	37.8 ± 0.5	7.7 ± 0.5
9	22422	β -D-allose	24.8 ± 0.1	26.9 ± 0.4	2.1 ± 0.4
10	24422	β -D-altrose	29.1 ± 0.2	28.7 ± 0.3	-0.5 ± 0.4
11	22442	β -D-gulose	25.2 ± 0.2	27.6 ± 0.4	2.4 ± 0.4
12	24442	β -D-idose	28.2 ± 0.1	32.7 ± 0.3	4.5 ± 0.3
13	42422	α -D-allose	26.9 ± 0.1	31.6 ± 0.3	4.7 ± 0.3
14	44422	α -D-altrose	29.4 ± 0.1	29.7 ± 0.4	0.3 ± 0.4
15	42442	α -D-gulose	24.6 ± 0.2	31.3 ± 0.3	6.7 ± 0.4
16	44442	α -D-idose	28.0 ± 0.1	32.6 ± 0.4	4.6 ± 0.4
17	22224	β -L-glucose	29.1 ± 0.1	30.8 ± 0.2	1.7 ± 0.2
18	24224	β -L-mannose	23.7 ± 0.1	30.1 ± 0.2	6.4 ± 0.2
19	22244	β -L-galactose	29.8 ± 0.1	28.0 ± 0.2	-1.7 ± 0.2
20	24244	β -L-talose	27.0 ± 0.1	30.4 ± 0.3	3.4 ± 0.3
21	42224	α -L-glucose	27.5 ± 0.2	31.2 ± 0.2	3.6 ± 0.3
22	44224	α -L-mannose	25.6 ± 0.2	27.2 ± 0.3	1.6 ± 0.4
23	42244	α -L-galactose	29.0 ± 0.2	27.0 ± 0.2	-2.0 ± 0.3
24	44244	α -L-talose	26.1 ± 0.2	27.0 ± 0.4	0.9 ± 0.4
25	22424	β -L-allose	31.2 ± 0.2	35.2 ± 0.8	4.0 ± 0.8
26	24424	β -L-altrose	23.6 ± 0.1	34.6 ± 0.9	10.9 ± 0.9
27	22444	β -L-gulose	29.5 ± 0.1	32.4 ± 0.3	3.0 ± 0.3
28	24444	β -L-idose	20.2 ± 0.1	33.2 ± 0.8	12.9 ± 0.8
29	42424	α -L-allose	33.6 ± 0.2	45.6 ± 1.4	12.1 ± 1.4
30	44424	α -L-altrose	33.8 ± 0.4	36.2 ± 1.4	2.4 ± 1.5
31	42444	α -L-gulose	30.8 ± 0.3	38.6 ± 0.9	7.8 ± 0.9
32	44444	α -L-idose	27.4 ± 0.2	33.5 ± 1.0	6.1 ± 1.0

Table 1 shows the 32 simulated stereoisomers, their 5-digit ID code, and the corresponding sugar names. For systems SUG_{vac}^{ab} and SUG_{water}^{ab} , the relative free energies of individual stereoisomers with respect to the reference state in kJ/mol are listed. The range in relative free energies amounts to 13.6 kJ/mol ($20.2 - 33.8$ kJ/mol) for system SUG_{vac}^{ab} and to 18.7 kJ/mol ($26.9 - 45.6$ kJ/mol) for system SUG_{water}^{ab} . In achiral environments such as vacuum and water, no differences in the

relative free energies are expected between enantiomers. In **Table 1**, sugar-pairs with codes 1 and 32, 2 and 31, 3 and 30, etc. represent enantiomers. Except for enantiomers β -D-talose (40.8 kJ/mol) and α -L-allose (45.6 kJ/mol) in system $SUG_{\text{water}}^{\text{ab}}$, the relative free energies for the enantiomers match very well within both systems: absolute differences between 0.1 – 1.8 kJ/mol in system $SUG_{\text{vac}}^{\text{ab}}$ and 0.1 – 2.6 kJ/mol in system $SUG_{\text{water}}^{\text{ab}}$ are roughly within the thermal noise of $k_{\text{B}}T$. The relative free energies for β -D-talose and α -L-allose match qualitatively as they are the two largest within system $SUG_{\text{water}}^{\text{ab}}$. The value for α -L-allose seems exceptionally high and omission of this value reduces the range for system $SUG_{\text{water}}^{\text{ab}}$ to 13.9 kJ/mol, similar to the vacuum value. Overall, the small free-energy differences between enantiomers give confidence in the applicability of the reference compound and in subsequent calculations. In the last column of **Table 1**, the calculated relative solvation free energies ($\Delta\Delta G_{\text{solv}}$) are listed. To the best of our knowledge, values for these quantities have not been reported in the literature previously, neither from experimental nor from computational sources. While the values of $\Delta\Delta G_{\text{solv}}$ are relative to the reference state, the differences between these values provide insights into the solvation of monosaccharides.

Hydrolysis at the pyranose C1 atom allows for interconversion between the α - and β -anomers characterized by a corresponding equilibrium. **Table 2** lists the free energy differences of β/α -anomers ($\Delta G_{\beta-\alpha}$) of all simulated pyranose D-stereoisomers in water, for which comparison with the available experimental data is possible. The calculated values were obtained from equation (3), while the experimental values were calculated from previously published experimental estimates of the β/α -pyranose ratios. The experimental values (at 30°C) were found to be largely temperature-insensitive [54], and can be readily compared to the simulation data obtained at 300 K or 26.85°C. Estimates of $\Delta G_{\beta-\alpha}$ from the experimental β/α -pyranose ratios were calculated according to

$$\Delta G_{\beta-\alpha} = -k_{\text{B}}T \ln\left(\frac{\beta}{\alpha}\right) \quad (10)$$

Table 2: Free energy differences of β/α -anomers of all simulated pyranose D-stereoisomers in kJ/mol for which experimental data is available. The values were obtained as differences of relative free energies of the two anomers calculated from the 100 ns simulation of system $SUG_{\text{water}}^{\text{ab}}$ or from previously published experimental values [54].

Sugar	$\Delta G_{\beta-\alpha}$ [kJ/mol]	
	Experiment [54]	Simulation
D-glucose	-1.2	-2.8
D-mannose	1.8	4.1
D-galactose	-1.7	-1.7
D-talose	1.0	3.0
D-allose	-4.1	-4.7
D-altrose	-1.1	-1.0
D-gulose	-4.8	-3.7
D-idose	-0.3	0.1

The $\Delta G_{\beta-\alpha}$ in **Table 2** obtained from MD simulations or experiment have very small absolute deviations in a range between 0.0 – 2.3 kJ/mol, which is smaller than the thermal noise, with a mean absolute deviation of 0.5 ± 1.3 kJ/mol.

For each of the 32 simulated monosaccharides of system $\text{SUG}_{\text{water}}^{\text{ab}}$, we investigated the occurrence of each of the 14 possible ring conformations [55,56] of the six-membered pyranose ring by correlating the observed ring conformations with the values of the improper dihedral angles in simulation $\text{SUG}_{\text{water}}^{\text{ab}}$. We found that sugars with code 1–16 (D-stereoisomers) occurred predominantly in the ${}^4\text{C}_1$ chair conformation and sugars with code 17–32 (L-stereoisomers) in the ${}^1\text{C}_4$ chair conformation. This again agrees with experimentally observed ring-conformational preferences of the D- or L-series of the studied aldohexopyranoses [55–57], which gives additional confidence in the conducted MD simulations.

4.3.2 Monosaccharide binding to PDH

Figure 4 shows (i) the occurrence of each of the 32 stereoisomers as a function of time and (ii) the number of occurrences with a lifetime ≥ 1 ps. The left two panels are derived from the 100 ns MD simulation of system $\text{SUG}_{\text{water}}^{\text{ab}}$, the right two panels represent the 50 ns MD simulation of system PDH- $\text{SUG}_{\text{md1}}^{\text{abc}}$ (pose A), which was selected as a representative example. System $\text{SUG}_{\text{water}}^{\text{ab}}$ nicely samples all stereoisomers and indicates many transitions between the monosaccharides, leading to good statistics for subsequent analysis. System PDH- $\text{SUG}_{\text{md1}}^{\text{abc}}$ (pose A) shows significantly less sampling and transitions of the stereoisomers. Therefore, six MD simulations (systems PDH- SUG^{a} , PDH- SUG^{ab} , PDH- SUG^{abc} ; two independent runs each) were conducted for each pose as mentioned previously. In some of the simulations of the PDH-SUG complexes, the unphysical reference state compound was observed to leave the active site. This may very well represent the proper behavior of these molecules, but unbound mixtures of PDH and SUG are (i) not expected to be relevant for real molecules binding to PDH and (ii) not part of the thermodynamic cycle to calculate the binding free energies according to equations (4) and (5). For this reason, simulations PDH- $\text{SUG}_{\text{md2}}^{\text{ab}}$ and PDH- $\text{SUG}_{\text{md2}}^{\text{abc}}$ for pose A and PDH- $\text{SUG}_{\text{md2}}^{\text{a}}$ and PDH- $\text{SUG}_{\text{md2}}^{\text{ab}}$ for pose B were excluded from the following analyses and four independent simulations of each pose remained. For the time series of relevant distances between PDH and SUG for all simulations see **Figures S1 and S2** in the supplementary material. The remaining four MD simulations for each pose were exponentially averaged according to equation (9) to calculate the free-energy differences (ΔG) of individual stereoisomers and their relative binding free energies ($\Delta\Delta G_{\text{bind}}$). According to **Figure 4**, system PDH- $\text{SUG}_{\text{md1}}^{\text{abc}}$ (pose A) clearly samples L-stereoisomers (sugar code 17-32; 5th digit of improper dihedral code is 4) better than D-stereoisomers (sugar code 1-16; 5th digit of improper dihedral code is 2). This is not surprising, as the transitions of the large $\text{CH}_2\text{-OH}$ group attached at this position are sterically the most hindered (see **Figure 2**).

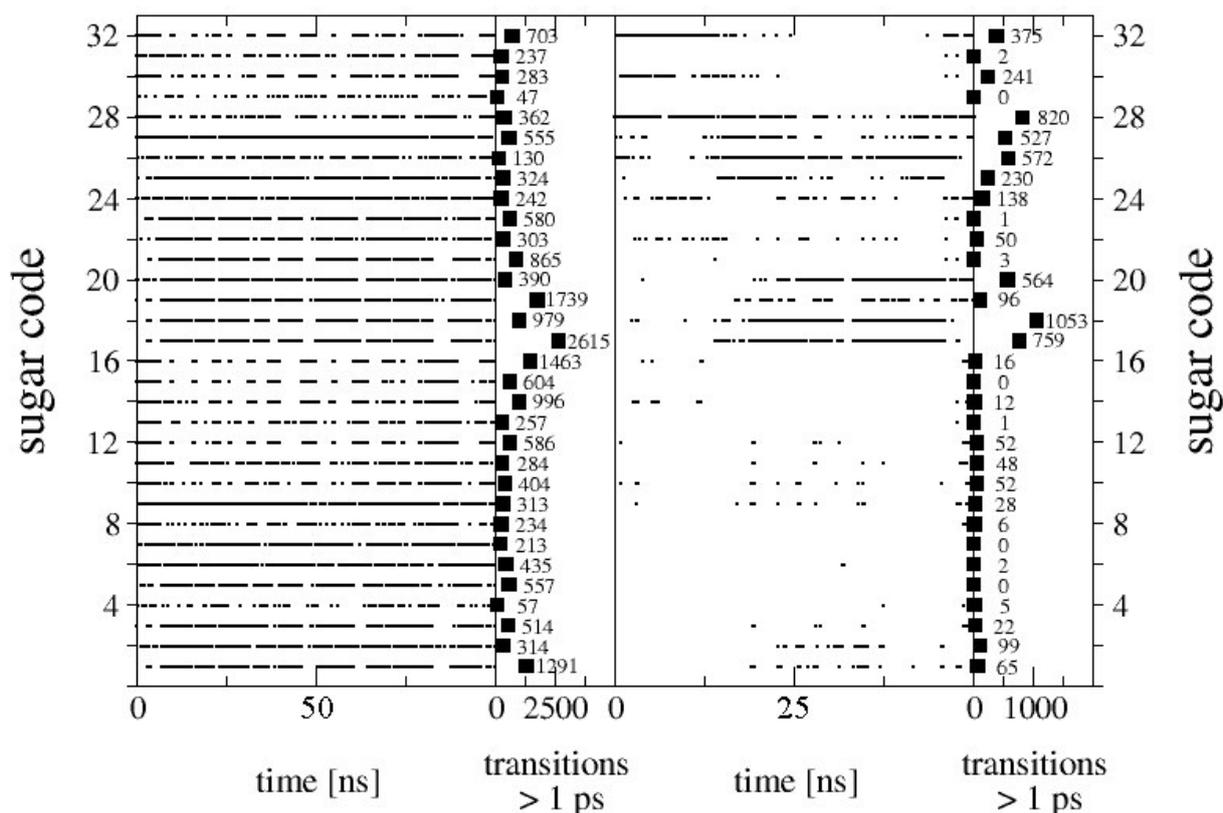


Figure 4: Indicated are (i) the occurrences of the 32 stereoisomers as a function of time and (ii) the number of occurrences with a lifetime ≥ 1 ps for a particular stereoisomer. The left two panels are derived from the 100 ns MD simulation of system SUG_{water}^{ab} , the right two panels represent the 50 ns MD simulation of system $PDH-SUG_{md1}^{abc}$ (pose A), which was selected as a representative example for the PDH-SUG complex.

Figure 5 shows the distributions of all five improper dihedrals (ID) for the MD simulations in water and in protein. For system SUG_{water}^{ab} (top five panels in **Figure 5**), the distributions of the ID are derived from the single 100 ns MD simulation, which nicely sampled all 32 possible stereoisomers (see **Figure 4**, left two panels). For system PDH-SUG (pose A) (middle five panels in **Figure 5**), and for system PDH-SUG (pose B) (lowest five panels in **Figure 5**), the occurrences of the IDs of the four selected MD simulations were arithmetically averaged. Except for ID5 in pose B and to a lesser extent ID3 in pose A, all improper dihedrals show fairly equal distributions and consequently very good sampling. In spite of lower occurrences for one configuration, ID5 (pose B) and ID3 (pose A) sample both stereoconfigurations. As mentioned previously, a large CH_2-OH group is attached at ID5 (compare **Figure 2**) and consequently transitions of this group are most sterically hindered in the MD simulations within the protein.

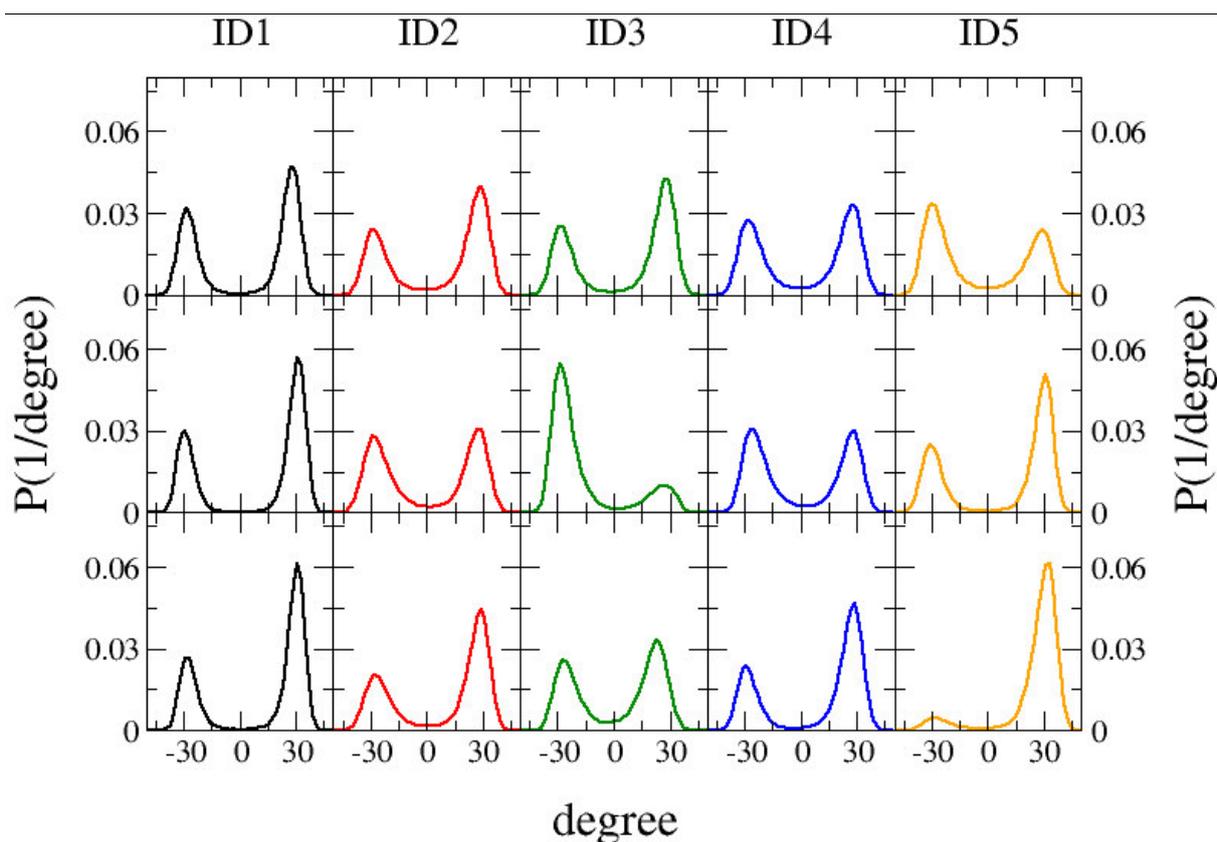


Figure 5: Distributions of all five improper dihedrals (ID) for the MD simulations in water and in protein. For system SUG_{water}^{ab} (top five panels), the distributions of the ID are derived from the single 100 ns MD simulation, which sampled all 32 possible stereoisomers (also compare to **Figure 4**, left two panels). For system PDH-SUG (pose A) (middle five panels), and for system PDH-SUG (pose B) (lowest five panels), the occurrences as observed in the four selected MD simulations were arithmetically averaged. Coloring scheme according to **Figure 2A**.

Table 3 lists the free-energy differences (ΔG) of the 32 stereoisomers simulated in system PDH-SUG (pose A or pose B). The reported ΔG values were obtained by exponentially averaging the four selected MD simulations for pose A or pose B. Because of the chiral environment within the protein, the span of ΔG values significantly increased (18.9 – 101.8 kJ/mol for pose A; 21.7 – 52.8 kJ/mol for pose B) compared to the MD simulation of system SUG_{water}^{ab} (26.9 – 45.6 kJ/mol; see **Table 1**). Moreover, the chiral protein-environment causes significant differences in ΔG between enantiomers. Enantiomers correspond to sugar-pairs with codes 1 and 32, 2 and 31, 3 and 30, etc. The ΔG between enantiomers range from 0.7 – 60.1 kJ/mol for pose A and from 1.8 – 24.1 kJ/mol for pose B (in absolute values). In addition, the relative binding free energies ($\Delta\Delta G_{bind}$) are listed in **Table 3**. They were calculated by subtracting the ΔG values for a certain monosaccharide in the MD simulation of system SUG_{water}^{ab} (see **Table 1**) from the ΔG of the identical monosaccharide in system PDH-SUG in either pose A or pose B (**Table 3**). Note that these values are relative to the reference states and only differences between them have physical relevance. The range for $\Delta\Delta G_{bind}$ for pose A is -11.7 to 56.2 kJ/mol and -10.1 to 23.1 kJ/mol for pose B. Interestingly, the $\Delta\Delta G_{bind}$

values for the α - and β -anomers of the D-stereoisomers of glucose are among the lowest of all simulated monosaccharides in both poses (range between -10 to -2.9 kJ/mol).

Table 3: Free-energy differences (ΔG) of the 32 stereoisomers simulated in system PDH-SUG, with SUG bound to PDH according to pose A or pose B. The reported ΔG were obtained by exponentially averaging the four selected MD simulations for each pose according to equation 9. The relative binding free energies ($\Delta\Delta G_{\text{bind}}$) were calculated by subtracting the ΔG in the MD simulation of system $\text{SUG}_{\text{water}}^{\text{ab}}$ (see **Table 1**) from the ΔG in system PDH-SUG in the corresponding pose (this table). The improper dihedral code has the same pattern as for **Table 1**.

Sugar code	Improper dihedral code	Sugar name	Pose A		Pose B	
			ΔG [kJ/mol]	$\Delta\Delta G_{\text{bind}}$ [kJ/mol]	ΔG [kJ/mol]	$\Delta\Delta G_{\text{bind}}$ [kJ/mol]
1	22222	β -D-glucose	22.5 ± 1.1	-9.1 ± 1.5	21.7 ± 0.5	-10.0 ± 1.1
2	24222	β -D-mannose	46.7 ± 2.1	8.7 ± 2.3	40.5 ± 2.1	2.5 ± 2.4
3	22242	β -D-galactose	29.2 ± 2.2	-5.3 ± 2.6	28.7 ± 1.1	-5.8 ± 1.7
4	24242	β -D-talose	41.7 ± 2.3	0.8 ± 3.4	39.9 ± 1.2	-1.0 ± 2.8
5	42222	α -D-glucose	31.5 ± 2.3	-2.9 ± 2.6	24.9 ± 1.5	-9.5 ± 1.8
6	44222	α -D-mannose	35.9 ± 1.3	2.1 ± 1.4	28.9 ± 2.3	-4.9 ± 2.4
7	42242	α -D-galactose	33.3 ± 1.9	-2.9 ± 2.3	28.7 ± 1.8	-7.5 ± 2.2
8	44242	α -D-talose	34.5 ± 1.4	-3.3 ± 1.5	30.9 ± 2.1	-6.9 ± 2.2
9	22422	β -D-allose	18.9 ± 0.6	-8.0 ± 0.7	25.7 ± 1.0	-1.1 ± 1.1
10	24422	β -D-altrose	36.6 ± 1.2	8.0 ± 1.2	27.8 ± 1.6	-0.9 ± 1.7
11	22442	β -D-gulose	26.6 ± 1.0	-1.0 ± 1.1	26.6 ± 0.5	-1.0 ± 0.7
12	24442	β -D-idose	40.0 ± 1.0	7.3 ± 1.1	32.5 ± 0.8	-0.2 ± 0.9
13	42422	α -D-allose	35.8 ± 1.0	4.2 ± 1.1	30.6 ± 1.9	-1.0 ± 1.9
14	44422	α -D-altrose	38.2 ± 1.8	8.6 ± 1.8	35.5 ± 1.9	5.8 ± 1.9
15	42442	α -D-gulose	36.0 ± 0.4	4.7 ± 0.5	35.0 ± 1.3	3.7 ± 1.3
16	44442	α -D-idose	36.3 ± 1.0	3.7 ± 1.1	32.4 ± 1.9	-0.2 ± 2.0
17	22224	β -L-glucose	40.1 ± 2.4	9.3 ± 2.4	39.5 ± 2.4	8.7 ± 2.4
18	24224	β -L-mannose	35.3 ± 0.4	5.2 ± 0.5	29.4 ± 2.7	-0.6 ± 2.7
19	22244	β -L-galactose	44.3 ± 1.8	16.3 ± 1.9	45.3 ± 2.1	17.3 ± 2.1
20	24244	β -L-talose	29.9 ± 1.2	-0.5 ± 1.2	38.9 ± 2.2	8.5 ± 2.3
21	42224	α -L-glucose	62.0 ± 2.5	30.8 ± 2.5	51.1 ± 2.5	20.0 ± 2.5
22	44224	α -L-mannose	39.1 ± 1.8	12.0 ± 1.8	40.4 ± 1.6	13.3 ± 1.7
23	42244	α -L-galactose	60.2 ± 2.5	33.2 ± 2.6	50.1 ± 1.2	23.1 ± 1.2
24	44244	α -L-talose	22.6 ± 1.8	-4.3 ± 1.9	40.0 ± 1.9	13.0 ± 1.9
25	22424	β -L-allose	30.6 ± 1.6	-4.6 ± 1.8	52.5 ± 1.5	17.3 ± 1.7
26	24424	β -L-altrose	24.7 ± 0.7	-9.9 ± 1.1	40.2 ± 2.8	5.7 ± 3.0
27	22444	β -L-gulose	37.2 ± 0.9	4.7 ± 0.9	36.2 ± 1.7	3.8 ± 1.8
28	24444	β -L-idose	24.9 ± 0.6	-8.3 ± 1.0	23.1 ± 1.5	-10.1 ± 1.7
29	42424	α -L-allose	101.8 ± 2.5	56.2 ± 2.9	47.3 ± 2.0	1.7 ± 2.4
30	44424	α -L-altrose	33.8 ± 1.1	-2.4 ± 1.8	52.8 ± 1.7	16.5 ± 2.3
31	42444	α -L-gulose	83.1 ± 1.4	44.5 ± 1.7	42.5 ± 1.9	3.9 ± 2.1
32	44444	α -L-idose	21.8 ± 1.7	-11.7 ± 2.0	31.7 ± 1.1	-1.9 ± 1.5

Table 4 gives an overview of the relative binding free energies ($\Delta\Delta G_{\text{bind}}$) calculated from the experimentally derived K_m values [23] and from the combined MD simulations of system PDH-SUG for either pose A or pose B. The experimental values were approximated from the corresponding K_m values according to the following formula:

$$\Delta G_{\text{bind}} = k_B T \ln K_m \quad (11)$$

Experimental data were available only for the four listed D-stereoisomers. Because the α - and β -anomers spontaneously interconvert in solution via mutarotation, they cannot be distinguished in experimental binding. The $\Delta\Delta G_{\text{bind}}$ values for pose A or pose B were calculated from the MD simulations by first exponentially averaging the free-energy differences of the α - and β -anomers of the respective D-stereoisomers simulated in system $\text{SUG}_{\text{water}}^{\text{ab}}$ (**Table 1**) or in system PDH-SUG in pose A or pose B (**Table 3**). Then, the averaged ΔG values in system $\text{SUG}_{\text{water}}^{\text{ab}}$ were subtracted from system PDH-SUG (pose A or pose B) to obtain the $\Delta\Delta G_{\text{bind}}$. For easier comparison, the $\Delta\Delta G_{\text{bind}}$ for D-glucose was set to zero. The $\Delta\Delta G_{\text{bind}}$ for pose A and pose B were not averaged, as the preference of the reference states for a certain pose is unknown. The $\Delta\Delta G_{\text{bind}}$ values derived from simulations of pose A agree well with experiment. Only the difference for D-talose between the experimental and calculated $\Delta\Delta G_{\text{bind}}$ (5.9 kJ/mol) is larger than the thermal noise. The agreement for pose B matches qualitatively, with differences between the experimental and calculated $\Delta\Delta G_{\text{bind}}$ values of 6.9 kJ/mol for D-mannose and 7.9 kJ/mol for D-talose, which are both above the thermal noise. To conclude, the $\Delta\Delta G_{\text{bind}}$ values derived from simulations and experiments match quite well.

Table 4: Comparison of the relative binding free energies ($\Delta\Delta G_{\text{bind}}$) calculated from experiments or MD simulations. The experimental values were estimated from the corresponding K_m values according to the formula $\Delta G_{\text{bind}} = k_B T \ln K_m$. All $\Delta\Delta G_{\text{bind}}$ values are reported relative to D-glucose, which was set to zero. Moreover, the $\Delta\Delta G_{\text{bind}}$ values for pose A and pose B were not averaged, because the preference of the reference states for a certain pose is unknown.

Sugar name	Michaelis constant K_m [mM] [23]	$\Delta\Delta G_{\text{bind}}$ [kJ/mol]		
		Experiment [23]	Simulation	
			Pose A	Pose B
D-glucose	0.82	0.0	0.0	0.0
D-mannose	108	12.3	10.9	5.4
D-galactose	1.05	0.6	3.7	3.4
D-talose	79.1	11.5	5.6	3.6

For successful oxidation, a hydride transfer takes place from the SUG-oxidation site to the N5 atom of FAD [13]. **Figure 6** shows the distances between H-atoms HC1–HC4 of SUG and the N5-atom of FAD. The position of the hydrogen atom in our united-atom representation of the reference state was determined according to ideal geometries and a C-H bond length of 0.1 nm. The occurrence of the distances of all four simulations of pose A and pose B were arithmetically

averaged. As reported previously [13], a 0.3 nm cutoff was considered in order for a hydride transfer to occur between HC1–HC4 and N5. Color codes are the same as in **Figure 2A**. In pose A (left panel), only HC1 (green) and HC2 (yellow) are below the mentioned cutoff. In pose B, only HC3 (red) and HC4 (blue) are below the 0.3 nm cutoff. This corresponds very well to previously published data for D-glucose oxidation by PDH, where pose A represents the C2 oxidation mode and pose B the C3 oxidation mode of this particular sugar [13].

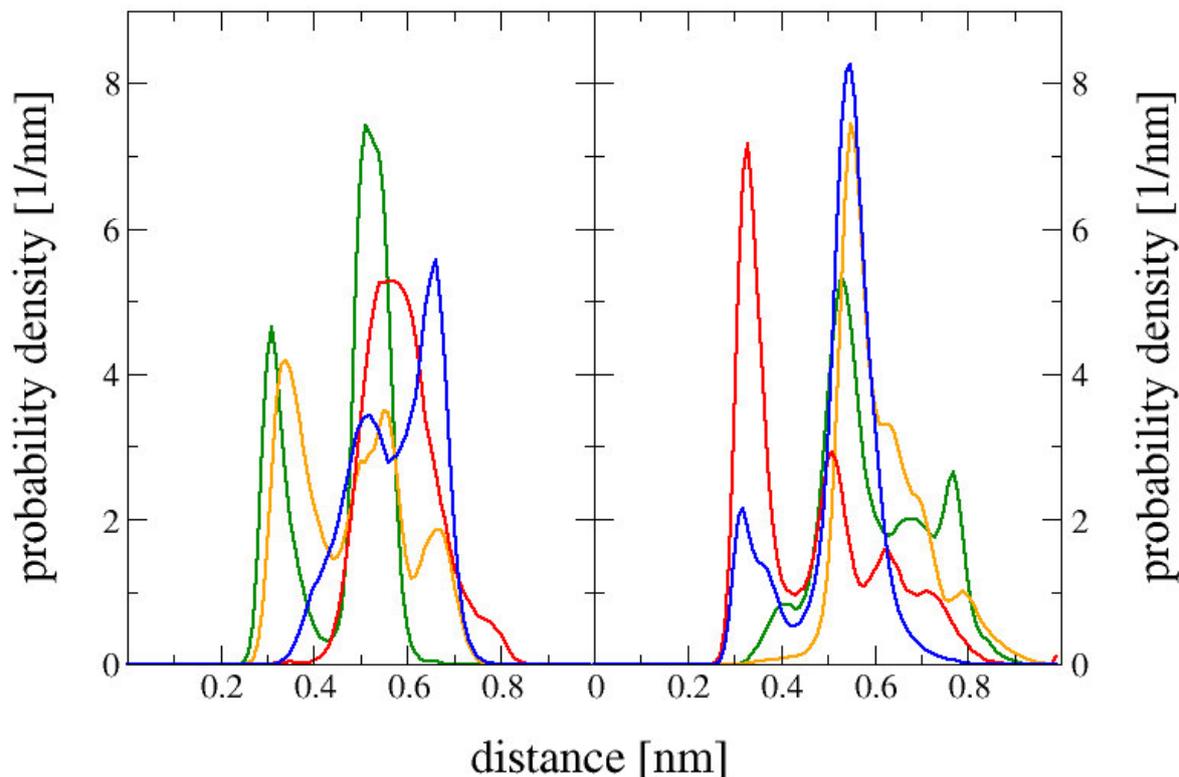


Figure 6: Distances between H-atoms HC1–HC4 of SUG and the N5-atom of FAD. The distances of all four simulations of pose A or pose B of system PDH-SUG were used. Left panel: pose A, right panel: pose B. Color codes match with **Figure 2A**: HC1 (green), HC2 (yellow), HC3 (red), and HC4 (blue).

In **Figure 7**, the average number of observations of distances between hydrogens attached to C1–C4 and the N5-atom in FAD below 0.3 nm for pose A and pose B are shown as calculated for all monosaccharides using equations 7 and 6. The bars in this logarithmic diagram are non-additive. These reactive distances are compared to the experimentally detected oxidation products. When the distances are below 0.3 nm, we will use the $\Delta\Delta G_{\text{bind}}$ to evaluate the likelihood of the corresponding monosaccharide to bind to PDH and consequently for a reaction to take place. Note that also low values of $\langle N(< 0.3 \text{ nm}) \rangle$ can already explain reactions, as substrate binding can easily be a much slower process than the actual reaction. The $\Delta\Delta G_{\text{bind}}$ value of β -D-glucose is set to zero in each pose and the $\Delta\Delta G_{\text{bind}}$ values of the other sugars are reported here relative to β -D-glucose.

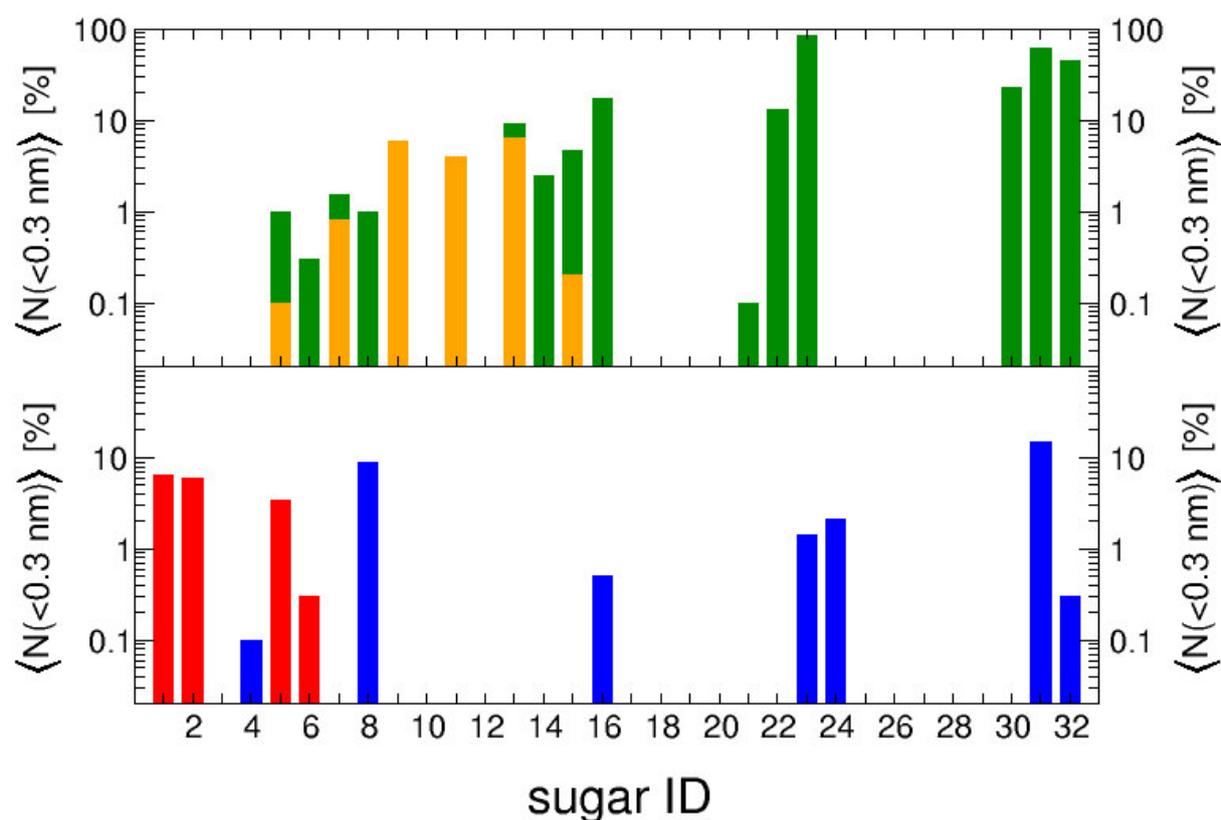


Figure 7: Reactive distances below 0.3 nm between hydrogens attached to C1–C4 of SUG and the N5-atom of FAD for pose A (upper panel) and pose B (lower panel). Color codes match with **Figure 2A**: HC1 (green), HC2 (yellow), HC3 (red), and HC4 (blue). The bars consisting of more than one color are non-additive.

Experimental D-glucose conversions yield (di-)oxidations at C2 and C3 [25]. This observation is reproduced by MD simulations (**Figure 7**): in pose A, 0.1% of the HC2-N5 distance is below 0.3 nm in α -D-glucose; in pose B, the HC3-N5 distance is below the chosen cutoff for 3.4% (α -D-glucose) and 6.5% (β -D-glucose). Again, this observation corresponds very well to our previously published work [13], where D-glucose is oxidized at C2 in pose A and at C3 in pose B. Experimentally, L-glucose is observed to have C2- and C3 (di-)oxidations as well. However, in our MD simulations we do not see any of the relevant HC2-N5 or HC3-N5 distances below the 0.3 nm cutoff. Moreover, its relative activity was experimentally measured to be 42% of D-glucose [25], which does not correspond to the predicted highly unfavorable $\Delta\Delta G_{\text{bind}}$ values between 18.5 – 39.9 kJ/mol for L-glucose bound in either pose. For D-galactose, MD simulations gave a HC2-N5 distance below 0.3 nm 0.8% of the time for α -D-galactose in pose A, which corresponds to its experimentally observed C2 oxidation [25]. Moreover, we predict a relatively favorable $\Delta\Delta G_{\text{bind}}$ value of 6.2 kJ/mol for α -D-galactose in pose A. The sugar D-mannose is a substrate for PDH, however, its oxidation sites have not yet been determined experimentally. The most prominent reactive distance for this sugar is HC3-N5 (β -D-mannose bound in pose B), which is below 0.3 nm for 5.9% of the time and has a predicted $\Delta\Delta G_{\text{bind}}$ value of 12.5 kJ/mol. For D-allose, C1 oxidation has been experimentally reported [25]. In the MD simulations, the corresponding HC1-N5 distance is below the 0.3 nm cutoff 2.9% of the time for α -D-allose (pose A). The predicted $\Delta\Delta G_{\text{bind}}$ value

for α -D-allose (pose A) is 13.3 kJ/mol, corresponding to the experimentally determined relative activity of 15% of D-glucose [25]. Experiments revealed solely C1 oxidation for D-talose, which was reproduced by MD simulations with the HC1-N5 distance below 0.3 nm 1.0% of the time for α -D-talose bound to PDH in pose A. The predicted $\Delta\Delta G_{\text{bind}}$ value for α -D-talose (pose A) is 5.8 kJ/mol, which agrees qualitatively and to a lesser extent quantitatively with the experimentally determined binding affinity according to its K_m value (see also **Table 4**). The HC4-N5 distance for α -D-talose (pose B) is below the 0.3 nm cutoff 8.8% of the time and the corresponding $\Delta\Delta G_{\text{bind}}$ value is reasonably low for binding (3.1 kJ/mol). Nevertheless, C4 oxidation is not reported experimentally, which might be caused by steric clashes of the adjacent hydroxymethyl-group attached to the C5 carbon resulting in poor binding. The last experimentally determined oxidation site for a monosaccharide investigated in this study is available for D-gulose, which is oxidized at C1 [25] only. In the MD simulations, the HC1-N5 distance for α -D-gulose (pose A) is indeed below 0.3 nm 4.4% of the time. The activity of D-gulose was reported to be 7% of D-glucose [25], which corresponds to an unfavorable $\Delta\Delta G_{\text{bind}}$ value for α -D-gulose (pose A) of 13.9 kJ/mol.

In addition to the experimentally determined oxidation sites, we made some striking observations during our distance analyses, which can direct future experiments. High percentages of reactive poses suggesting C1 oxidation are observed for the following sugars bound in pose A: α -D-idose (17.2%), α -L-mannose (13.3%), α -L-galactose (85.5%), α -L-altrose (22.9%), α -L-gulose (62.3%), and α -L-idose (45.0%). Some of these possible oxidation products can be neglected, as the predicted $\Delta\Delta G_{\text{bind}}$ value for the corresponding monosaccharides is very unfavorable in pose A: α -L-gulose (53.6 kJ/mol), α -L-galactose (42.3 kJ/mol), and α -L-mannose (21.1 kJ/mol). Others could have low, but measurable activities: α -D-idose (12.8 kJ/mol) and α -L-altrose (6.7 kJ/mol). Lastly, oxidation for α -L-idose (-2.6 kJ/mol) with HC1-N5 below 0.3 nm 45% of the time is predicted in pose A. For pose B, we predict HC4 oxidation of α -L-gulose, for which the HC4-N5 distance is below 0.3 nm 14.9% of the time and the predicted $\Delta\Delta G_{\text{bind}}$ value is 13.9 kJ/mol, allowing for low but measurable activity.

Interestingly, monosaccharides bound to PDH in pose A, for which additional oxidations were observed, are all α -compounds and oxidized at HC1. This can be rationalized, as the hydroxyl-group attached to C1 defines whether a sugar is an α - or β -anomer. Consequently, if the HC1-N5 is within the reactive distance, the hydroxyl-group attached to that C1 has to be on the opposite side of the HC1, which (in pose A) corresponds to the α -anomer of the respective sugars.

4.4 Conclusion

In this study, we presented a generalized approach to simulate monosaccharide solvation in water, as well as binding and product formation in the enzyme PDH. Introducing changes to the monosaccharide topology according to **Figure 2** created systems SUG^a , SUG^{ab} , and SUG^{abc} , out of which system SUG^{ab} was selected as the most suitable reference state for subsequent analysis in water. This allowed for sampling of all 32 possible aldohexopyranoses in only one MD simulation of the reference compound in water or using a handful of simulations of the reference state compounds within PDH.

Free energy calculations according to the one-step perturbation method revealed that systems $\text{SUG}_{\text{vac}}^{\text{ab}}$ and $\text{SUG}_{\text{water}}^{\text{ab}}$ show a similar range of relative free energies for the simulated monosaccharides. Moreover, the relative free energies for the enantiomer-pairs (sugar codes 1 and 32, 2 and 31, etc.) match very well within systems $\text{SUG}_{\text{vac}}^{\text{ab}}$ and $\text{SUG}_{\text{water}}^{\text{ab}}$. Because both vacuum and water represent an achiral environment, this outcome is expected and gives confidence in the conducted simulations. We reported calculated values for the relative solvation free energies ($\Delta\Delta G_{\text{solv}}$) of all 32 aldohexopyranoses (**Table 1**). To the best of our knowledge, these $\Delta\Delta G_{\text{solv}}$ values have not been reported previously, giving new fundamental insights into the solvation of aldohexopyranoses. For all simulated pyranose D-stereoisomers, we report the free energy differences of the corresponding β/α -anomers in water ($\Delta G_{\beta-\alpha}$). The deviations from experimental values [23,25] are very small (within a range of 0.0 – 2.3 kJ/mol), further increasing the confidence in the conducted simulations. In addition, the pyranose ring conformations for each of the 32 stereoisomers were investigated. Sugars with codes 1–16 (D-stereoisomers) occurred predominantly in the ${}^4\text{C}_1$ chair conformation and sugars with codes 17–32 (L-stereoisomers) in the ${}^1\text{C}_4$ chair conformation. These results are in line with the experimentally obtained preferences of ring conformations for D- or L-stereoisomers [55–57], furthermore strengthening the validity of the performed simulations.

For system PDH-SUG, six MD simulations of 50 ns each were performed for either pose A or pose B. Of these six, two were discarded for each pose from subsequent analysis, because the monosaccharide left the active site. For each pose, the results of the four selected MD simulations were exponentially averaged. All 32 possible stereoisomers were sampled extensively in pose A and pose B. Compared to system $\text{SUG}_{\text{water}}^{\text{ab}}$ (26.9 – 45.6 kJ/mol), the chiral protein environment caused a significant increase in the span of the free-energy differences (ΔG): 18.9 – 101.8 kJ/mol for pose A and 21.7 – 52.8 kJ/mol for pose B. The relative binding free energies ($\Delta\Delta G_{\text{bind}}$) range between -11.7 to 56.2 kJ/mol for pose A and between -10.1 to 23.1 kJ/mol for pose B. Where available, the calculated $\Delta\Delta G_{\text{bind}}$ values were compared to the experimental binding free energies estimated from K_m values. For pose A, the agreement is quite reasonable. Pose B shows qualitative agreement between calculated and experimental $\Delta\Delta G_{\text{bind}}$ values. Taking an arbitrary cutoff of +15 kJ/mol relative to β -D-glucose for binding, the predicted binding affinities indicate that measurable binding

free energies can be expected for 21 monosaccharides in pose A as well as in pose B. This is in line with the expected promiscuity of the PDH enzyme and suggests that these monosaccharides can be anticipated to inhibit the enzyme and possibly are also substrates themselves.

A distance analysis between hydrogens attached to the sugar carbons 1–4 and the N5 atom of FAD revealed that monosaccharide oxidation is possible at HC1 or HC2 in pose A and at HC3 or HC4 in pose B, which is in line with previously published findings [13]. We could reproduce the experimentally detected oxidation products by monitoring the HCX–N5 distance for D-glucose, D-galactose, D-allose, D-talose, and D-gulose. Only for L-glucose, the experimentally observed C2- or C3- oxidation could not be reproduced by the HCX–N5 distance analysis. With a combination of HCX–N5 distance analysis and binding free energy calculations, we predict oxidation products for some sugars, which have not yet been reported experimentally: low but measurable oxidation at HC1 for L-altrose and D-idose as well as at HC3 for D-mannose and at HC4 for L-gulose; strong oxidation at HC1 for L-idose – a challenge for future experiments.

To conclude, this study presents a generalized approach to simulate all 32 possible aldohexopyranoses in the course of just a few simulations. It contributes to the rationalization of PDH's substrate promiscuity with a combination of binding free energies and distance analyses for each sugar. This provides insights into PDH's applicability in bioelectrochemistry. We believe that this approach is readily transferable to other promiscuous enzymes, whose substrates differ mainly in the stereochemistry of their reactive groups.

4.5 References

- 1 Hult K & Berglund P (2007) Enzyme promiscuity: mechanism and applications. *Trends Biotechnol.* **25**, 231–238.
- 2 Arevalo JH, Taussig MJ & Wilson IA (1993) Molecular basis of crossreactivity and the limits of antibody-antigen complementarity. *Nature* **365**, 859–863.
- 3 Mariuzza RA (2006) Multiple paths to multispecificity. *Immunity* **24**, 359–361.
- 4 James LC & Tawfik DS (2003) Conformational diversity and protein evolution – a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **28**, 361–368.
- 5 Jensen RA (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425.
- 6 Tawfik OK and DS (2010) Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505.
- 7 Wackett LP (2004) Evolution of enzymes for the metabolism of new chemical inputs into the environment. *J. Biol. Chem.* **279**, 41259–41262.
- 8 Miller BG & Raines RT (2004) Identifying latent enzyme activities: Substrate ambiguity within modern bacterial sugar kinases. *Biochemistry* **43**, 6387–6392.
- 9 Kim J & Copley SD (2007) Why metabolic enzymes are essential or nonessential for growth of *Escherichia coli* K12 on glucose. *Biochemistry* **46**, 12501–12511.
- 10 Kurakin A (2007) Self-organization versus watchmaker: Ambiguity of molecular recognition and design charts of cellular circuitry. *J. Mol. Recognit.* **20**, 205–214.
- 11 Nobeli I, Favia AD & Thornton JM (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol.* **27**, 157–167.
- 12 Khersonsky O, Roodveldt C & Tawfik DS (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* **10**, 498–508.
- 13 Graf MMH, Bren U, Haltrich D & Oostenbrink C (2013) Molecular dynamics simulations give insight into D-glucose dioxidation at C2 and C3 by *Agaricus meleagris* pyranose dehydrogenase. *J. Comput. Aided Mol. Des.* **27**, 295–304.
- 14 Basdevant N, Weinstein H & Ceruso M (2006) Thermodynamic basis for promiscuity and selectivity in protein–protein interactions: PDZ domains, a case study. *J. Am. Chem. Soc.* **128**, 12766–12777.
- 15 Oostenbrink C, de Ruiter A, Hritz J & Vermeulen N (2012) Malleability and versatility of cytochrome P450 active sites studied by molecular simulations. *Curr. Drug Metab.* **13**, 190–196.
- 16 Bornscheuer UT & Kazlauskas RJ (2004) Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed.* **43**, 6032–6040.
- 17 Copley SD (2003) Enzymes with extra talents: Moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol.* **7**, 265–272.
- 18 Davies GJ, Gloster TM & Henrissat B (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr. Opin. Struct. Biol.* **15**, 637–645.

- 19 Cavener DR (1992) GMC oxidoreductases: A newly defined family of homologous proteins with diverse catalytic activities. *J. Mol. Biol.* **223**, 811–814.
- 20 Volc J, Kubátová E, Wood DA & Daniel G (1997) Pyranose 2-dehydrogenase, a novel sugar oxidoreductase from the basidiomycete fungus *Agaricus bisporus*. *Arch. Microbiol.* **167**, 119–125.
- 21 Volc J, Kubátová E, Daniel G, Sedmera P & Haltrich D (2001) Screening of basidiomycete fungi for the quinone-dependent sugar C-2/C-3 oxidoreductase, pyranose dehydrogenase, and properties of the enzyme from *Macrolepiota rhacodes*. *Arch. Microbiol.* **176**, 178–186.
- 22 Kujawa M, Volc J, Halada P, Sedmera P, Divne C, Sygmund C, Leitner C, Peterbauer C & Haltrich D (2007) Properties of pyranose dehydrogenase purified from the litter-degrading fungus *Agaricus xanthoderma*. *FEBS J.* **274**, 879–894.
- 23 Sygmund C, Kittl R, Volc J, Halada P, Kubátová E, Haltrich D & Peterbauer CK (2008) Characterization of pyranose dehydrogenase from *Agaricus meleagris* and its application in the C-2 specific conversion of D-galactose. *J. Biotechnol.* **133**, 334–342.
- 24 Krondorfer I, Lipp K, Brugger D, Staudigl P, Sygmund C, Haltrich D & Peterbauer CK (2014) Engineering of pyranose dehydrogenase for increased oxygen reactivity. *PLoS ONE* **9**, e91145.
- 25 Peterbauer CK & Volc J (2010) Pyranose dehydrogenases: Biochemical features and perspectives of technological applications. *Appl. Microbiol. Biotechnol.* **85**, 837–848.
- 26 Tan TC, Spadiut O, Wongnate T, Sucharitakul J, Krondorfer I, Sygmund C, Haltrich D, Chaiyen P, Peterbauer CK & Divne C (2013) The 1.6 Å crystal structure of pyranose dehydrogenase from *Agaricus meleagris* rationalizes substrate specificity and reveals a flavin intermediate. *PLoS ONE* **8**, e53567.
- 27 Sedmera P, Halada P, Kubátová E, Haltrich D, Přikrylová V & Volc J (2006) New biotransformations of some reducing sugars to the corresponding (di)dehydro(glycosyl) aldoses or aldonic acids using fungal pyranose dehydrogenase. *J. Mol. Catal. B Enzym.* **41**, 32–42.
- 28 Oostenbrink C (2012) Free energy calculations from one-step perturbations. In *Computational drug discovery and design*. Baron, R., Ed. Humana Press, New York; Vol. 819, pp 487–499.
- 29 Liu H, Mark AE & van Gunsteren WF (1996) Estimating the relative free energy of different molecular states with respect to a single reference state. *J. Phys. Chem.* **100**, 9485–9494.
- 30 Oostenbrink C & van Gunsteren WF (2004) Free energies of binding of polychlorinated biphenyls to the estrogen receptor from a single simulation. *Proteins Struct. Funct. Bioinforma.* **54**, 237–246.
- 31 De Beer SBA, Venkataraman H, Geerke DP, Oostenbrink C & Vermeulen NPE (2012) Free energy calculations give insight into the stereoselective hydroxylation of α -ionones by engineered Cytochrome P450 BM3 mutants. *J. Chem. Inf. Model.* **52**, 2139–2148.
- 32 Lai B, Nagy G, Garate JA & Oostenbrink C (2014) Entropic and enthalpic contributions to stereospecific ligand binding from enhanced sampling methods. *J. Chem. Inf. Model.* **54**, 151–158.

- 33 Hritz J & Oostenbrink C (2009) Efficient free energy calculations for compounds with multiple stable conformations separated by high energy barriers. *J. Phys. Chem. B* **113**, 12711–12720.
- 34 Garate JA & Oostenbrink C (2013) Free-energy differences between states with different conformational ensembles. *J. Comput. Chem.* **34**, 1398–1408.
- 35 Lin Z, Kornfeld J, Mächler M & van Gunsteren WF (2010) Prediction of folding equilibria of differently substituted peptides using one-step perturbation. *J. Am. Chem. Soc.* **132**, 7276–7278.
- 36 Lin Z, Oostenbrink C & Gunsteren WF van (2014) On the use of one-step perturbation to investigate the dependence of NOE-derived atom–atom distance bound violations of peptides upon a variation of force-field parameters. *Eur. Biophys. J.* **43**, 113–119.
- 37 Tan T-C, Haltrich D & Divne C (2011) Regioselective control of β -D-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy. *J. Mol. Biol.* **409**, 588–600.
- 38 Kujawa M, Ebner H, Leitner C, Hallberg BM, Prongjit M, Sucharitakul J, Ludwig R, Rudsander U, Peterbauer C, Chaiyen P, Haltrich D & Divne C (2006) Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J. Biol. Chem.* **281**, 35104–35115.
- 39 Schmid N, Christ CD, Christen M, Eichenberger AP & van Gunsteren WF (2012) Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **183**, 890–903.
- 40 Oostenbrink C, Villa A, Mark AE & Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676.
- 41 Lins RD & Hünenberger PH (2005) A new GROMOS force field for hexopyranose-based carbohydrates. *J. Comput. Chem.* **26**, 1400–1412.
- 42 Berendsen HJC, Postma JPM, Van Gunsteren WF & Hermans J (1981) Interaction models for water in relation to protein hydration. In *Intermolecular Forces*. Reidel, Dordrecht. pp 331–342.
- 43 Amadei A, Chillemi G, Ceruso MA, Grottesi A & Di Nola A (2000) Molecular dynamics simulations with constrained roto-translational motions: Theoretical basis and statistical mechanical consistency. *J. Chem. Phys.* **112**, 9–23.
- 44 Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
- 45 Ryckaert J-P, Ciccotti G & Berendsen HJ. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.
- 46 Miyamoto S & Kollman PA (1992) Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962.
- 47 Tironi IG, Sperb R, Smith PE & van Gunsteren WF (1995) A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459.

- 48 Heinz TN, van Gunsteren WF & Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **115**, 1125–1136.
- 49 Zwanzig RW (1954) High-temperature equation of state by a perturbation method. I. nonpolar gases. *J. Chem. Phys.* **22**, 1420–1426.
- 50 Torrie GM & Valleau JP (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem. Phys. Lett.* **28**, 578–581.
- 51 Torrie GM & Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23**, 187–199.
- 52 De Beer SBA, Glättli A, Hutzler J, Vermeulen NPE & Oostenbrink C (2011) Molecular dynamics simulations and free energy calculations on the enzyme 4-hydroxyphenylpyruvate dioxygenase. *J. Comput. Chem.* **32**, 2160–2169.
- 53 Chodera JD, Swope WC, Pitera JW, Seok C & Dill KA (2007) Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J. Chem. Theory Comput.* **3**, 26–41.
- 54 Sinnott ML (2007) Carbohydrate chemistry and biochemistry. In *Carbohydrate Chemistry and Biochemistry*. Royal Society of Chemistry, Cambridge. pp. 1–4.
- 55 Rao VSR (1998) Conformation of Carbohydrates. CRC Press. p11.
- 56 Hansen HS & Hünenberger PH (2010) Using the local elevation method to construct optimized umbrella sampling potentials: Calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water. *J. Comput. Chem.* **31**, 1–23.
- 57 Gabius H-J (2009) The Sugar Code: Fundamentals of Glycosciences. Weinheim: Wiley-VCH. p 17.

4.6 Supplementary material

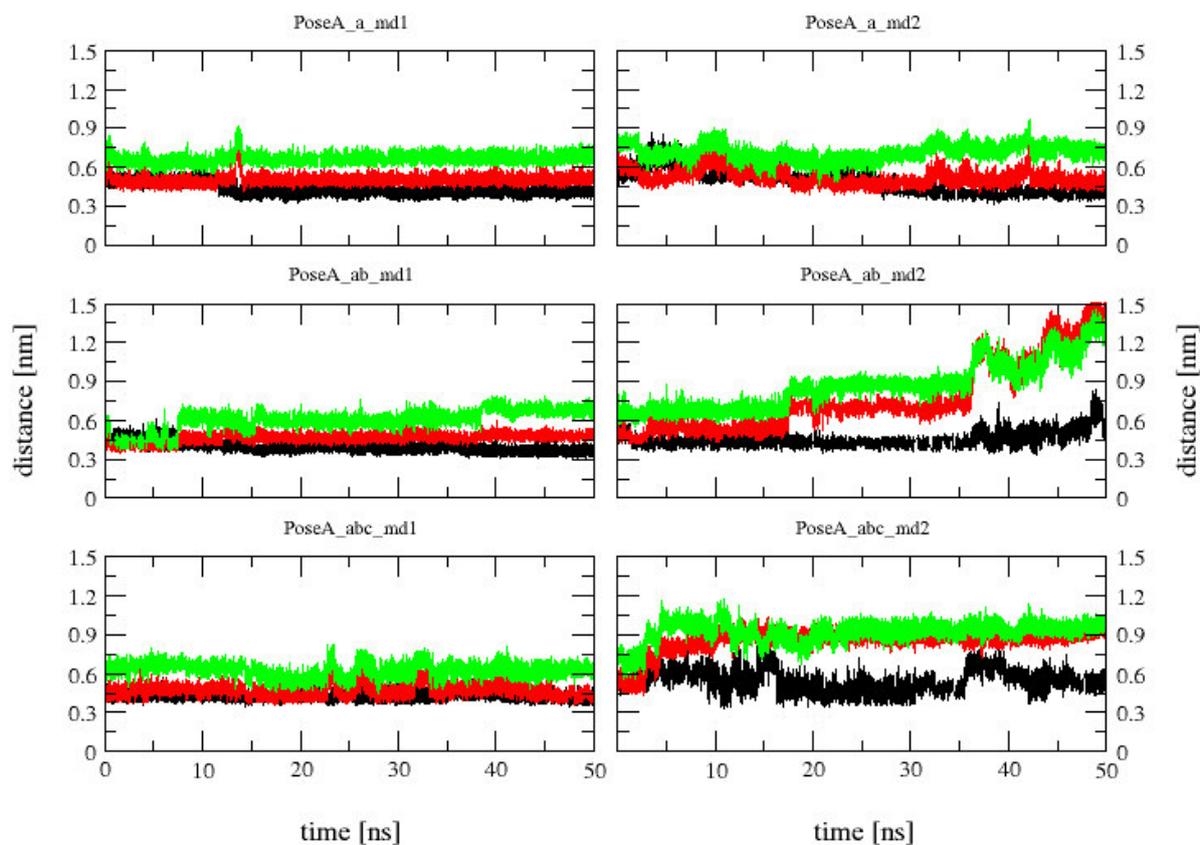


Figure S1: Distances between selected atoms of PDH and SUG in system PDH-SUG (pose A) used to monitor whether SUG left the binding site. In the upper two panels, the SUG-topology was altered according to SUG^a , in the middle two panels according to SUG^{ab} , and in the lowest two panels according to SUG^{abc} (see **Figure 2** and text for more details). The first column represents the first repeat of the MD simulations for each of the changed SUG-topologies (md1) and the second column the second repeat (md2). Colors indicate the following distances: Glu-392(NE2)-SUG(C4) (black), Val-511(C)-SUG(C1) (red), and Val-511(N)-SUG(C5) (green). For pose A, the MD simulations of systems PDH-SUG $_{md2}^{ab}$ and PDH-SUG $_{md2}^{abc}$ were discarded, because the SUG left PDH's active site.

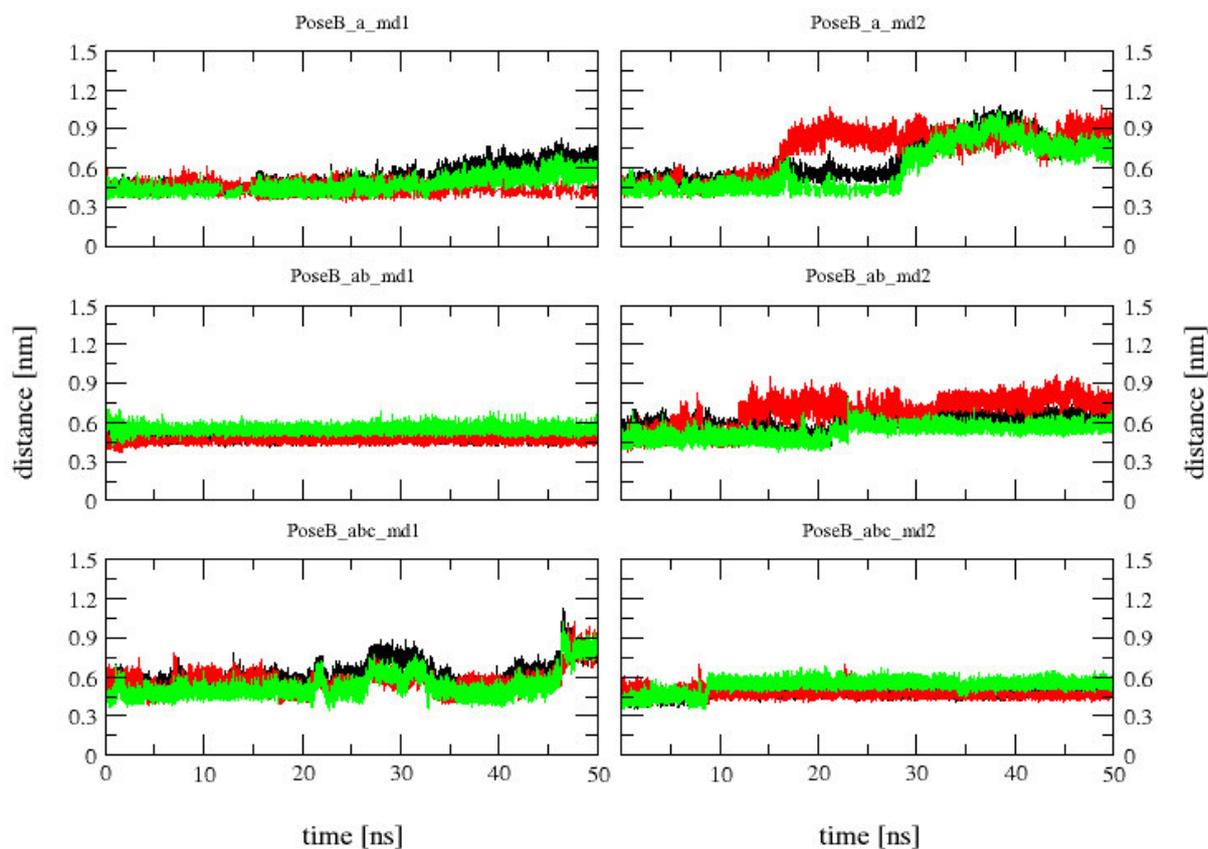


Figure S2: Distances between selected atoms of PDH and SUG in system PDH-SUG (pose B) used to monitor whether SUG left the binding site. In the upper two panels, the SUG-topology was altered according to SUG^a , in the middle two panels according to SUG^{ab} , and in the lowest two panels according to SUG^{abc} (see **Figure 2** and text for more details). The first column represents the first repeat of the MD simulations for each of the changed SUG-topologies (md1) and the second column the second repeat (md2). Colors indicate the following distances: Val-511(C)-SUG(C3) (black), Gln-392(CD)-SUG(C1) (red), and Val-511(N)-SUG(C4) (green). For pose B, the MD simulations of systems PDH-SUG $_{md2}^a$ and PDH-SUG $_{md2}^{ab}$ were discarded, because the SUG left PDH's active site. Although SUG left the active site in system PDH-SUG $_{md1}^{abc}$ in the last 2-3 ns, it was still used for subsequent analysis.


```

# CHB: harmonic force constant
# B0: bond length at minimum energy
#
#      CB          CHB          B0
1.57000e+07    3.14000e+05    1.00000e-01
1.87000e+07    3.74000e+05    1.00000e-01
1.23000e+07    2.92273e+05    1.09000e-01
3.70000e+07    9.28256e+05    1.12000e-01
1.66000e+07    5.02283e+05    1.23000e-01
1.34000e+07    4.18750e+05    1.25000e-01
1.20000e+07    4.18176e+05    1.32000e-01
8.87000e+06    3.13803e+05    1.33000e-01
1.06000e+07    3.75007e+05    1.33000e-01
1.18000e+07    4.17460e+05    1.33000e-01
# 10
1.05000e+07    3.77076e+05    1.34000e-01
1.17000e+07    4.20170e+05    1.34000e-01
1.02000e+07    3.77318e+05    1.36000e-01
1.10000e+07    4.18968e+05    1.38000e-01
8.66000e+06    3.34640e+05    1.39000e-01
1.08000e+07    4.17334e+05    1.39000e-01
8.54000e+06    3.34768e+05    1.40000e-01
8.18000e+06    3.34546e+05    1.43000e-01
9.21000e+06    3.76671e+05    1.43000e-01
6.10000e+06    2.51225e+05    1.43500e-01
# 20
8.71000e+06    3.76429e+05    1.47000e-01
5.73000e+06    2.51020e+05    1.48000e-01
7.64000e+06    3.34693e+05    1.48000e-01
8.60000e+06    3.76749e+05    1.48000e-01
8.37000e+06    3.76650e+05    1.50000e-01
5.43000e+06    2.50909e+05    1.52000e-01
7.15000e+06    3.34749e+05    1.53000e-01
4.84000e+06    2.50915e+05    1.61000e-01
4.72000e+06    2.50811e+05    1.63000e-01
2.72000e+06    1.72361e+05    1.78000e-01
# 30
5.94000e+06    3.76406e+05    1.78000e-01
5.62000e+06    3.76416e+05    1.83000e-01
3.59000e+06    2.51077e+05    1.87000e-01
6.40000e+05    5.01811e+04    1.98000e-01
6.28000e+05    5.02400e+04    2.00000e-01
5.03000e+06    4.18657e+05    2.04000e-01
5.40000e+05    5.27483e+04    2.21000e-01
2.32000e+07    4.64000e+05    1.00000e-01
1.21000e+07    2.92820e+05    1.10000e-01
8.12000e+06    5.01908e+05    1.75800e-01
# 40
8.04000e+06    3.76417e+05    1.53000e-01
4.95000e+06    3.71825e+05    1.93799e-01
8.10000e+06    5.01811e+05    1.76000e-01
1.31000e+07    4.19259e+05    1.26500e-01
1.03000e+07    3.75435e+05    1.35000e-01
8.71000e+06    4.64532e+05    1.63299e-01
2.68000e+06    2.93088e+05    2.33839e-01
2.98000e+06    5.02215e+05    2.90283e-01
2.39000e+06    3.73116e+05    2.79388e-01
2.19000e+06    3.71385e+05    2.91189e-01
# 50
3.97000e+06    3.42526e+05    2.07700e-01
3.04000e+06    5.02225e+05    2.87407e-01
END
BONDH
# NBNH: number of bonds involving H atoms in solute
5
# IBH, JBH: atom sequence numbers of atoms forming a bond
# ICBH: bond type code
# IBH JBH ICBH
# 1 2 1

```

```
      5      6      1
      8      9      1
     11     12      1
     16     17      1
END
BOND
# NBON: number of bonds NOT involving H atoms in solute
12
# IB, JB: atom sequence numbers of atoms forming a bond
# ICB: bond type code
#   IB      JB      ICB
      2       3      20
      3       4      26
      3      13      26
      4       5      20
      4       7      26
      7       8      20
      7      15      26
     10      11      20
     10      13      26
     13      14      20
# 10
     14      15      20
     15      16      20
END
BONDANGLEBENDTYPE
# NTTY: number of bond angle types
54
# CT: force constant (based on potential
#     harmonic in the angle cosine)
# CHT: force constant (based on potential
#       harmonic in the angle)
# T0: bond angle at minimum energy in degrees
#   CT          CHT          T0
     3.80000e+02    1.15501e-01    9.00000e+01
     4.20000e+02    1.27686e-01    9.00000e+01
     4.05000e+02    1.21771e-01    9.60000e+01
     4.75000e+02    1.40083e-01    1.00000e+02
     4.20000e+02    1.21222e-01    1.03000e+02
     4.90000e+02    1.40285e-01    1.04000e+02
     4.65000e+02    1.27888e-01    1.08000e+02
     2.85000e+02    7.69125e-02    1.09500e+02
     3.20000e+02    8.63861e-02    1.09500e+02
     3.80000e+02    1.02627e-01    1.09500e+02
# 10
     4.25000e+02    1.14807e-01    1.09500e+02
     4.50000e+02    1.21574e-01    1.09500e+02
     5.20000e+02    1.40521e-01    1.09500e+02
     4.50000e+02    1.21423e-01    1.09600e+02
     5.30000e+02    1.40487e-01    1.11000e+02
     5.45000e+02    1.40451e-01    1.13000e+02
     5.00000e+01    1.22966e-02    1.15000e+02
     4.60000e+02    1.14885e-01    1.15000e+02
     6.10000e+02    1.52417e-01    1.15000e+02
     4.65000e+02    1.14219e-01    1.16000e+02
# 20
     6.20000e+02    1.52361e-01    1.16000e+02
     6.35000e+02    1.53360e-01    1.17000e+02
     3.90000e+02    8.89104e-02    1.20000e+02
     4.45000e+02    1.01476e-01    1.20000e+02
     5.05000e+02    1.15184e-01    1.20000e+02
     5.30000e+02    1.20895e-01    1.20000e+02
     5.60000e+02    1.27749e-01    1.20000e+02
     6.70000e+02    1.52880e-01    1.20000e+02
     7.80000e+02    1.78011e-01    1.20000e+02
     6.85000e+02    1.53127e-01    1.21000e+02
# 30
     7.00000e+02    1.53174e-01    1.22000e+02
```

```

4.15000e+02      8.87438e-02      1.23000e+02
7.30000e+02      1.52669e-01      1.24000e+02
3.75000e+02      7.64902e-02      1.25000e+02
7.50000e+02      1.53141e-01      1.25000e+02
5.75000e+02      1.14487e-01      1.26000e+02
6.40000e+02      1.27447e-01      1.26000e+02
7.70000e+02      1.53365e-01      1.26000e+02
7.60000e+02      1.27755e-01      1.32000e+02
2.21500e+03      1.21127e-01      1.55000e+02
# 40
9.13500e+04      7.26402e-02      1.80000e+02
4.34000e+02      1.17243e-01      1.09500e+02
4.84000e+02      1.33765e-01      1.07570e+02
6.32000e+02      1.66891e-01      1.11300e+02
4.69000e+02      1.40245e-01      9.74000e+01
5.03000e+02      1.40260e-01      1.06750e+02
4.43000e+02      1.21084e-01      1.08530e+02
6.18000e+02      1.67047e-01      1.09500e+02
5.07000e+02      1.40086e-01      1.07600e+02
4.48000e+02      1.21033e-01      1.09500e+02
# 50
5.24000e+02      1.40180e-01      1.10300e+02
5.32000e+02      1.40257e-01      1.11400e+02
6.36000e+02      1.53054e-01      1.17200e+02
6.90000e+02      1.52948e-01      1.21400e+02
END
BONDANGLEH
# NTHEH: number of bond angles involving H atoms in solute
5
# ITH, JTH, KTH: atom sequence numbers
#   of atoms forming a bond angle in solute
# ICTH: bond angle type code
#   ITH   JTH   KTH ICTH
#     1     2     3   12
#     4     5     6   12
#     7     8     9   12
#    10    11    12   12
#    15    16    17   12
END
BONDANGLE #Modified according to Figure 2C
# NTHE: number of bond angles NOT
#   involving H atoms in solute
17
# IT, JT, KT: atom sequence numbers of atoms
#   forming a bond angle
# ICT: bond angle type code
#   IT   JT   KT   ICT
#     2   3   4   9 #Last value set to 8
#     2   3  13   9 #Last value set to 8
#     4   3  13   8
#     3   4   5   9 #Last value set to 8
#     3   4   7   8
#     5   4   7   9 #Last value set to 8
#     4   7   8   9 #Last value set to 8
#     4   7  15   8
#     8   7  15   9 #Last value set to 8
#    11  10  13   9
# 10
#     3  13  10   8
#     3  13  14   9 #Last value set to 8
#    10  13  14   9 #Last value set to 8
#    13  14  15  10 #Last value set to 8
#     7  15  14   9 #Last value set to 8
#     7  15  16   9 #Last value set to 8
#    14  15  16   9 #Last value set to 8
END
IMPDIHEDRALTYPE
# NQTY: number of improper dihedrals

```

```
4
# CQ: force constant of improper dihedral per degrees square
# Q0: improper dihedral angle at minimum energy in degrees
#           CQ           Q0
#   5.10000e-02    0.00000e+00
#   1.02000e-01    3.52644e+01
#   2.04000e-01    0.00000e+00
#   1.02000e-01   -3.52644e+01
END
IMPDIHEDRALH
# NQHIH: number of improper dihedrals
#         involving H atoms in the solute
0
# IQH,JQH,KQH,LQH: atom sequence numbers
#         of atoms forming an improper dihedral
# ICQH: improper dihedral type code
#   IQH   JQH   KQH   LQH ICQH
END
IMPDIHEDRAL #Modified according to Figure 2A
# NQHI: number of improper dihedrals NOT
#         involving H atoms in solute
5 #This value was set to 0
# IQ,JQ,KQ,LQ: atom sequence numbers of atoms
#         forming an improper dihedral
# ICQ: improper dihedral type code
#   IQ   JQ   KQ   LQ ICQ
#   3    10   14   13   2 #This line was deleted
#   4     5    7    3    2 #This line was deleted
#   7    14   16   15   2 #This line was deleted
#  13    2    4    3    2 #This line was deleted
#  15    4    8    7    2 #This line was deleted
END
TORSDIHEDRALTYPE
# NPTY: number of dihedral types
41
# CP: force constant
# PD: phase-shift angle
# NP: multiplicity
#           CP           PD   NP
#   2.67000  180.00000   1
#   3.41000  180.00000   1
#   4.97000  180.00000   1
#   5.86000  180.00000   1
#   9.35000  180.00000   1
#   9.45000  180.00000   1
#   2.79000   0.00000   1
#   5.35000   0.00000   1
#   1.53000  180.00000   2
#   5.86000  180.00000   2
#  10
#   7.11000  180.00000   2
#  16.70000  180.00000   2
#  24.00000  180.00000   2
#  33.50000  180.00000   2
#  41.80000  180.00000   2
#   0.00000   0.00000   2
#   0.41800   0.00000   2
#   2.09000   0.00000   2
#   3.14000   0.00000   2
#   5.09000   0.00000   2
#  20
#  16.70000   0.00000   2
#   1.05000   0.00000   3
#   1.26000   0.00000   3
#   1.30000   0.00000   3
#   2.53000   0.00000   3
#   2.93000   0.00000   3
#   3.19000   0.00000   3
```

```

3.65000    0.00000    3
3.77000    0.00000    3
3.90000    0.00000    3
# 30
4.18000    0.00000    3
4.69000    0.00000    3
5.44000    0.00000    3
5.92000    0.00000    3
7.69000    0.00000    3
8.62000    0.00000    3
9.50000    0.00000    3
0.00000    0.00000    4
1.00000    180.00000   6
1.00000    0.00000    6
# 40
3.77000    0.00000    6
END
DIHEDRALH
# NPHIH: number of dihedrals involving H atoms in solute
6
# IPH, JPH, KPH, LPH: atom sequence numbers
# of atoms forming a dihedral
# ICPH: dihedral type code
# IPH   JPH   KPH   LPH ICPH
# 1     2     3     4   30
# 7     4     5     6   30
# 15    7     8     9   30
# 13    10    11    12  30
# 14    15    16    17   2
# 14    15    16    17  32
END
DIHEDRAL #Modified according to Figure 2B
# NPHI: number of dihedrals NOT involving H atoms in solute
19
# IP, JP, KP, LP: atom sequence numbers
# of atoms forming a dihedral
# ICP: dihedral type code
# IP   JP   KP   LP   ICP
# 2    3    4    5   18 #Last value set to 17
# 2    3    4    7   17
# 13   3    4    5   17
# 13   3    4    7   34 #Last value set to 22
# 2    3    13   10  17
# 4    3    13   14  17
# 4    3    13   14  34 #Last value set to 22
# 3    4    7    8   17
# 3    4    7    15  34 #Last value set to 22
# 5    4    7    8   18 #Last value set to 17
# 10
# 5    4    7    15  17
# 4    7    15   14  17
# 4    7    15   14  34 #Last value set to 22
# 4    7    15   16  17
# 8    7    15   16  18 #Last value set to 17
# 11   10   13   14   5
# 11   10   13   14  37
# 3    13   14   15  29 #Last value set to 22
# 13   14   15    7  29 #Last value set to 22
END
CROSSDIHEDRALH
# NPHIH: number of cross dihedrals involving H atoms in solute
0
# APH, BPH, CPH, DPH, EPH, FPH, GPH, HPH: atom sequence numbers
# of atoms forming a dihedral
# ICCH: dihedral type code
# APH   BPH   CPH   DPH   EPH   FPH   GPH   HPH ICCH
END
CROSSDIHEDRAL

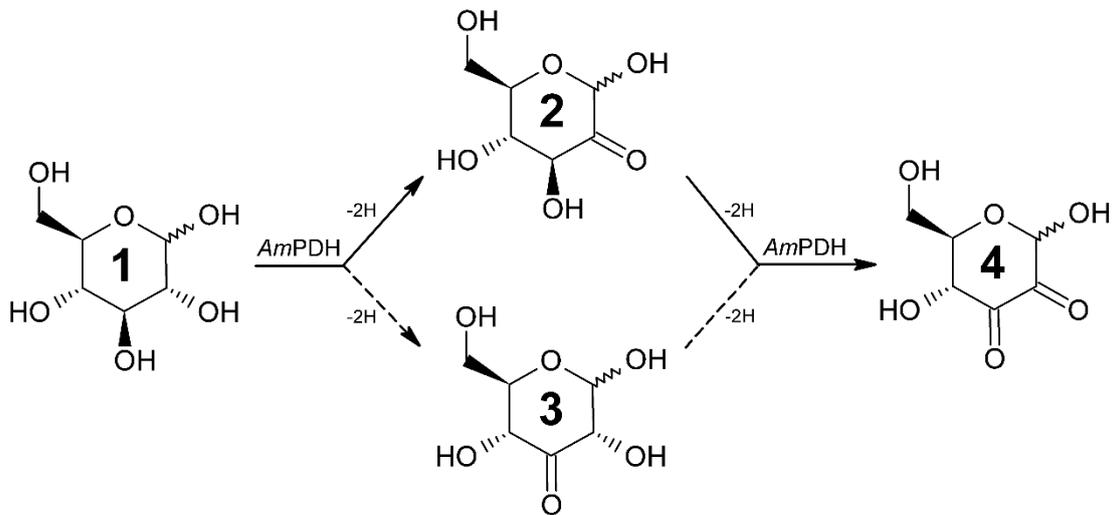
```

Ligand Promiscuity of PDH

```
# NPPC: number of cross dihedrals NOT involving H atoms in solute
0
# AP, BP, CP, DP, EP, FP, GP, HP: atom sequence numbers
#   of atoms forming a dihedral
# ICC: dihedral type code
#   AP   BP   CP   DP   EP   FP   GP   HP   ICC
END
```

CHAPTER 5

PDH – Experiments and Simulations



Abstract

Monomeric *Agaricus meleagris* pyranose dehydrogenase (*Am*PDH) belongs to the glucose–methanol–choline (GMC) family of oxidoreductases and carries a FAD cofactor, which is covalently tethered to His-103. *Am*PDH (di)oxidizes mono- and polysaccharides occurring during lignin degradation, making it interesting for sugar transformations in the food- and pharmaceutical industry or for bioelectrochemical applications. To date, little is known about structure–dynamics–function relationships of active site residues in *Am*PDH, hampering its rational engineering towards the desired applications. To tackle this problem, eight active site variants were generated, heterologously expressed in *Pichia pastoris*, characterized by biochemical, biophysical, and computational means, and compared to the wild type. The covalent linkage to FAD was disrupted only in variant H103A, which resulted in a 29 mV decrease in reduction potential. Steady-state kinetics revealed that His-512 is the sole catalytic base, while Gln-392, Tyr-510, Val-511, and His-556 are important for substrate binding. Molecular dynamics (MD) simulations and free energy calculations reproduced experimental K_m values and predicted D-glucose oxidation sites, which were validated by GC-MS measurements and culminated in an adapted D-glucose oxidation scheme for *Am*PDH. Stopped-flow analysis revealed that *Am*PDH equally prefers C2 and C3 oxidized substrates, which is consistent with its substrate promiscuity. For variant H556A, a decrease in the reductive half-reaction by almost three orders of magnitude and in reduction potential by an estimated 8 mV indicated the importance of His-556 for sugar oxidation. The oxygen reactivities for H556A and H103A were increased 3.7- and 4.1-fold, which was attributed to steric and electrochemical effects.

Authors

Michael Graf, Jeerus Sucharitakul, Urban Bren, Dinh Binh Chu, Gunda Koellensperger, Stephan Hann, Paul Furtmüller, Christian Obinger, Chris Oostenbrink, Pimchai Chaiyen & Dietmar Haltrich (2015) Active site residues in *Agaricus meleagris* pyranose dehydrogenase: Structure–dynamics–function relationships unraveled by biochemical, biophysical, and computational means. Manuscript in preparation.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF) Doctoral Program BioToP – Biomolecular Technology of Proteins (FWF W1224), by the Vienna Science and Technology Fund (WWTF LS08-QM03), by the European Research Council (ERC, 260408), the Slovenian Research Agency (ARRS; grant numbers P1-0002 and J1-5448) which is gratefully acknowledged. Dinh Binh Chu acknowledges Erasmus Mundus Action 2 (www.eurasia2.cz) for financial support. Funding for measuring the redox potential and pre-steady-state kinetics at Mahidol University in Thailand and measurements at the department of Chemistry at BOKU, Vienna is gratefully acknowledged.

5.1 Introduction

The oxidoreductase pyranose dehydrogenase from the litter decomposing fungus *Agaricus meleagris* (*Am*PDH; EC 1.1.99.29; PDB code 4H7U) is a glycoprotein with approximately 65 kDa carrying a monocovalently linked FAD cofactor [1]. The level of glycosylation depends on its source, ranging from about 7% (*Agaricus meleagris*) [2] to 30% (*Pichia pastoris*) [3]. Together with aryl-alcohol oxidase (AAO) [4], choline oxidase (CHO) [5], glucose 1-oxidase (GOX) [6], and pyranose 2-oxidase (P2O) [7], amongst others, *Am*PDH belongs to the structural family of glucose–methanol–choline (GMC) oxidoreductases [1,8]. PDH was initially identified in the fungi *Agaricus bisporus* [9] and *Macrolepiota rhacodes* [10], however, PDH from *Agaricus meleagris* [11] is best characterized to date. Although *Am*PDH can oxidize many different carbohydrates, D-glucose (GLC) is one of its favored substrates as judged by its catalytic efficiency [2]. Studies on the preferred oxidation sites revealed that *Am*PDH is able to (di)oxidize mono- and polysaccharides at C1–C4 [11–14], and molecular dynamics (MD) simulations suggested an oxidation mechanism for the preferred substrate D-glucose [15]. This extraordinarily broad substrate promiscuity as well as its molecular properties [2] make *Am*PDH interesting for sugar transformations in the food- and pharmaceutical industry as well as for bioelectrochemical applications [1].

Producing PDH in *Agaricus meleagris* takes approximately six weeks and yields only a few mg of pure enzyme [2]. In order to provide sufficient amounts within a short period of time several attempts to heterologously express *Am*PDH in *Aspergillus* spp. [16], *Escherichia coli*, and *Pichia pastoris* [3] were made. *P. pastoris* proved to be the heterologous expression host of choice as judged by the highest yields of intact enzyme and its short cultivation time of about one week. This paved the way to elucidate the crystal structure of *Am*PDH (PDB code 4H7U) [17] as well as to probe its oxygen reactivity [18] and the covalent FAD linkage [19]. With the available crystal structure, *in silico* studies were conducted to investigate the reaction mechanism [15] and binding free energies of all 32 possible aldohexopyranoses [20] by means of molecular dynamics (MD) simulations.

Previous characterizations of *Am*PDH have either focused on the wild-type *Am*PDH [2,17] or on identifying variants with increased oxygen reactivity employing a semi-rational approach coupled with high-throughput screening [18]. Only the variant with the highest oxygen reactivity identified in the latter study [18] was subject to a more detailed biochemical and biophysical analysis [19]. The current work represents the first comprehensive rational investigation of the most important active site amino acids in *Am*PDH (**Figure 1**) [15,17,18]. The aim of the study was to obtain a detailed understanding of structure-function relationships of the active site in *Am*PDH. This should ultimately direct novel approaches to better engineer *Am*PDH towards the desired applications. Moreover, the appreciation of these fundamental principles in *Am*PDH could possibly be extended to other related GMC oxidoreductases.

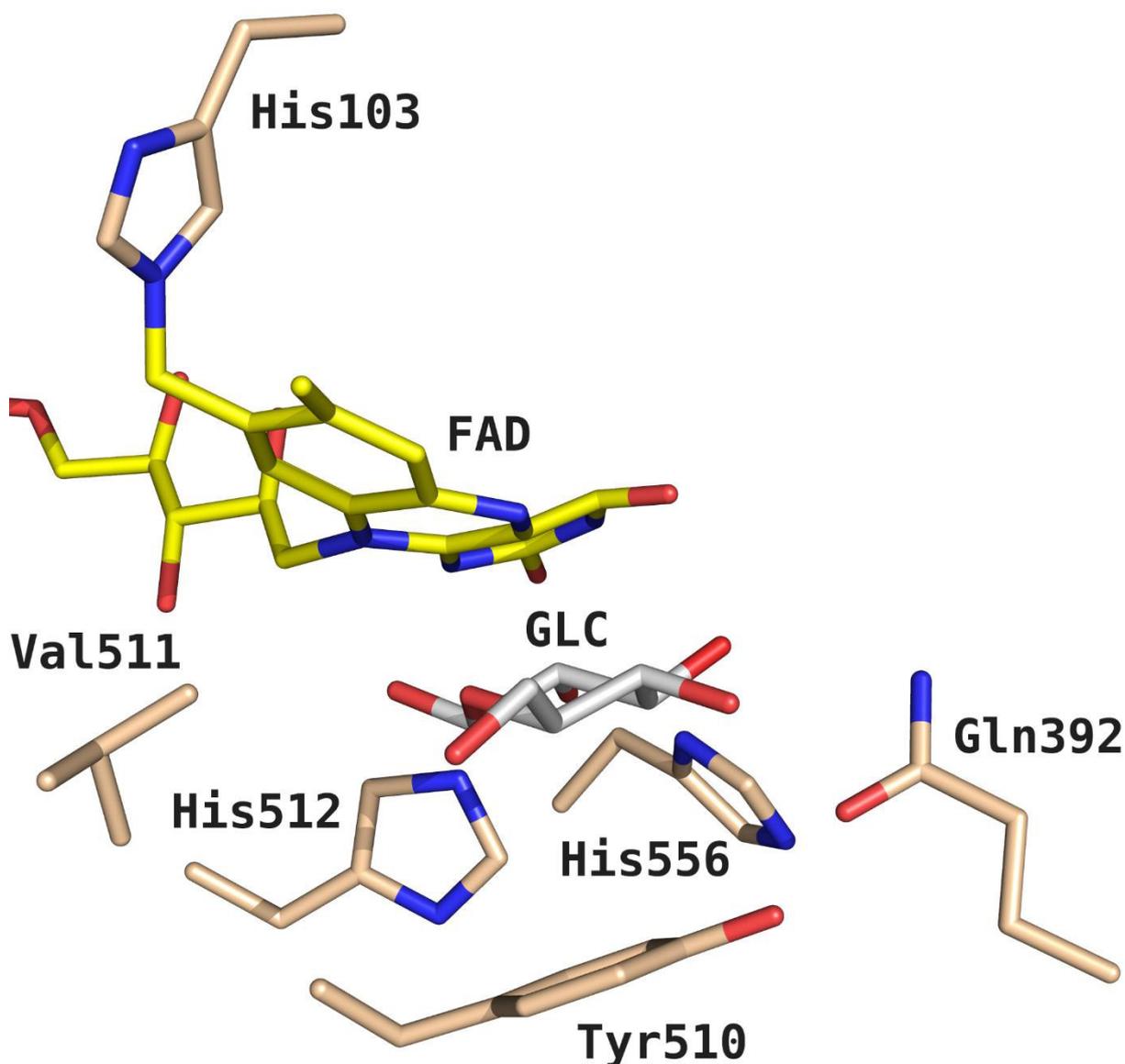


Figure 1: Active site of pyranose dehydrogenase from *Agaricus meleagris* (AmPDH; PDB code 4H7U) with bound β -D-glucose (GLC) according to pose A (see materials and methods for more information on binding poses). GLC coordinates from the closely related pyranose 2-oxidase (TmP2O) from *Trametes multicolor* (PDB code 3PL8) were grafted into AmPDH after superimposing the X-ray structures of both enzymes. Atom-coloring scheme: carbon (beige, protein; yellow, FAD; white, ligand), nitrogen (blue), and oxygen (red). The figure was generated using PyMOL (<http://www.pymol.org/>).

Previous studies on *AmPDH* suggested that both active site histidines (His-512 and His-556) potentially act as catalytic base [15,18] and that mutational studies on Gln-392, Tyr-510, and Val-511 would affect GLC binding and product formation [15]. To test these hypotheses, active site variants Q392A, Y510A, V511F, V511W, H512A, H556A, and H556N were generated by site-directed mutagenesis. Additionally, the covalent linkage between His-103 and FAD was disrupted in variant H103A to probe the influence of the covalent bond on the enzyme. The variants were heterologously expressed in *P. pastoris* as reported previously [3,18] and purified according to a

purification protocol that was adapted from Sygmond and coworkers [3] and utilizes a newly introduced hexahistidine-tag (His₆-tag).

Introducing mutations into a protein possibly disrupts the overall structure and decreases its stability [21]. Therefore, the secondary structure and thermostability of every variant was tested by electronic circular dichroism (ECD) spectroscopy [22] and *ThermoFAD* measurements [23], respectively, and compared to *Am*PDH. In the closely related P2O, mutational studies on amino acids not directly involved in the covalent attachment of FAD, still resulted in mixed populations of (non-)covalently bound flavin [24]. Consequently, *Am*PDH and all of its variants were tested for covalent flavinylation with TCA/acetone precipitation experiments [19,25]. It is known that covalently linked FAD has a significantly increased redox potential [26]. Consequently, the redox potential was determined for *Am*PDH as well as variant H103A and H556A – for the first time according to Massey’s method [27]. It is not exactly known to date, why flavoenzymes react with oxygen or not, however, even small changes – especially in the active site – can have significant effects [28–30]. Consequently, all generated active site variants were routinely tested for their oxygen reactivity with an Amplex Red/horseradish peroxidase based assay [18].

To probe whether both active site histidines in *Am*PDH can act as catalytic base [15,18] and if the other residues are involved in GLC binding [15], the apparent K_m and k_{cat} values were determined for the preferred substrate GLC. MD simulations and free energy calculations helped to explain the obtained K_m values. For the first time, we could predict GLC oxidation sites for *Am*PDH and its variants by MD simulations and validate them *via* GC-MS measurements. Finally, rapid reaction experiments elucidated the preference of *Am*PDH for C2 or C3 oxidized substrates as well as the effects of mutation H556A on both half-reactions.

5.2 Materials and Methods

5.2.1 Chemicals and vectors

All chemicals were of the highest available purity and purchased from Sigma Aldrich (St. Louis, MO, USA), VWR (Radnor, PA, USA) and Roth (Karlsruhe, Germany). Primers were ordered from LGC Genomics (Vienna, Austria) or Sigma Aldrich (Austria, Vienna). Restriction endonucleases, T4 DNA ligase, and Phusion polymerase were from Thermo Fisher Scientific Biosciences (St. Leon-Rot, Germany). GoTaq polymerase was purchased from Promega (Madison, WI, USA). Zeocin and the pPICZB vector were obtained from Invitrogen (Carlsbad, CA, USA). The 2-keto-GLC standard, meso-erythritol, LC-MS grade distilled water, ethoxylamine hydrochloride, water-free pyridine, and *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) in water-free pyridine were purchased from Sigma Aldrich (Vienna, Austria).

5.2.2 Strains and media

Strains and media used for this study were essentially the same as previously reported [18]. In short, *Escherichia coli* strain NEB5 α was obtained from New England Biolabs (Ipswich, MA, USA) and *Pichia pastoris* strain X33 from Invitrogen (Carlsbad, CA, USA). YPD plates contained 20 g/L peptone from casein, 10 g/L yeast extract, 4 g/L GLC, 15 g/L agar, and 100 mg/L Zeocin. LB low-salt (LB-LS) plates contained 10 g/L peptone from casein, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar, and 25 mg/L Zeocin. For fermentations, the basal salts medium with 4.35 mL/L PTM₁ trace salts was used, which is described in detail by Invitrogen or in [18].

5.2.3 Plasmid construction for expression in *P. pastoris*

A modified purification approach as reported previously [3,18] was adopted to purify active site variants, which involved affinity chromatography using a hexahistidine-tag (His₆-tag). When expressing the protein of interest with the original pPICZB plasmid from Invitrogen, a *myc*-epitope followed by a His₆-tag is added to the C-terminus of the protein. In order to alter the protein of interest as little as possible, the pPICZB plasmid was slightly modified for this study. When using the primer-pair pPICZB-6His-fw and pPICZB-6His-*Xba*I-rv (see **Table S1**) the *myc*-epitope was removed by following the *Dpn*I-method [31], creating pPICZB-His₆. The PCR-product was purified from an agarose gel, digested with *Dpn*I for 2 h at 37°C for degradation of methylated template-DNA. After another purification step, the pPICZB vector void of the *myc*-epitope was transformed into chemically competent *E. coli* NEB5 α . Positive transformants were selected on LB-LS plates containing Zeocin and subsequent sequencing. The confirmed plasmid was amplified in *E. coli* NEB5 α , purified, and stored at -20°C.

The *Agaricus meleagris* *pdh1* (*ampdh*) gene in the pPICZB vector as utilized by Krondorfer et al. [18] was amplified with the primer-pair *AmpDH-NotI-fw* and *AmpDH-XbaI-rv* (see **Table S1**), thereby introducing a 3'-*NotI*- and a 5'-*XbaI*-restriction site. The purified product was digested with the endonucleases *NotI* and *XbaI*, and ligated into the equally treated pPICZB-His₆ vector, yielding pPICZB-His₆-*AmpDH*, which was transformed into chemically competent *E. coli* NEB5 α . Positive transformants were selected on LB-LS plates containing the selection marker Zeocin and subsequent sequencing. An *E. coli* NEB5 α colony containing the verified plasmid was proliferated, the pPICZB-His₆-*AmpDH* purified and finally stored at -20°C.

The active site mutations H103A, Q392A, Y510A, V511F, V511W, H512A, H556A, and H556N were each introduced into a separate pPICZB-His₆-*AmpDH* vector by site-directed mutagenesis as described earlier in this section. The overlapping primer pairs for introducing the respective mutation were H103A-fw/H103A-rv, Q392A-fw/Q392A-rv, Y510A-fw/Y510A-rv, V511F-fw/V511-rv, V511W-fw/V511-rv, H512A-fw/H512A-rv, H556A-fw/H556A-rv, H556N-fw/H556N-rv (see **Table S1**). The mutations were verified by sequencing. Plasmids with the correct mutation were linearized with *PmeI* at 37°C for 2 h, purified, and transformed into electro-competent *P. pastoris* X33. Positive transformants were selected on YPD plates containing Zeocin. Additionally, plasmid integration was tested by colony PCR, for which the universal primers 5'-AOX1 and 3'-AOX1 were utilized.

5.2.4 Protein expression and purification

The wild-type *Agaricus meleagris* PDH1 (*AmpDH*) and the H556A variant were produced as described in [3] but in a 42-L computer-controlled stirred tank reactor (Applikon, Schiedam, The Netherlands) with an initial volume of 20 L basal salts fermentation medium. The variants H103A, Q392A, Y510A, V511F, V511W, H512A, and H556N were produced as described previously [18] in a 7-L computer-controlled fermenter (MBR, Wetzikon, Switzerland) with an initial volume of 4 L basal salts fermentation medium.

AmpDH and H556A were purified at room temperature according to a 4-step purification scheme (**Table 1**): first, the volume of the fermentation supernatant was reduced by a factor of approximately 10 by cross-flow filtration (Microza ultrafiltration module, Pall Austria Filter GmbH, Vienna, Austria) through a membrane with a 10-kDa cutoff.

The concentrate was subsequently purified by immobilized metal affinity chromatography (IMAC) with a Ni²⁺-charged Chelating Sepharose Fast Flow column (65 mL; flow rate 10 mL min⁻¹; GE Healthcare, Little Chalfont, UK), equilibrated with 100 mM potassium phosphate buffer pH 7.0, 1 M NaCl, and 5 mM imidazole. After a washing step of 3 column volumes (CV), *AmpDH* or H556A were eluted with a linear gradient from 5 to 500 mM imidazole. Fractions with PDH activity were pooled. This step is depicted as IMAC I in **Table 1**.

(NH₄)₂SO₄ was added to the pooled IMAC I-fractions to 40% saturation and loaded on a Phenyl Sepharose Fast Flow column (200 mL; flow rate 10 mL min⁻¹; GE Healthcare) equilibrated with 50 mM potassium phosphate buffer (pH 6.5, 40% saturation (NH₄)₂SO₄). After washing the column with 3 CV of the same buffer, *Am*PDH or H556A were eluted with a linear gradient of starting buffer to 50 mM potassium phosphate buffer (pH 6.5) in 1 CV. Fractions with PDH activity were pooled. This step is depicted as HIC in **Table 1**.

As a polishing step, a second IMAC purification (IMAC II in **Table 1**) was conducted: the pooled HIC-fractions were loaded on 1-4 in series connected HisTrap HP columns (5 mL each; flow rate 2 mL min⁻¹, GE Healthcare) equilibrated as for IMAC I. For washing and elution, the same condition as for IMAC I were applied. Fractions with PDH activity were pooled.

The variants H103A, Q392A, Y510A, V511F, V511W, H512A, and H556N were purified with a combination of HIC and IMAC II (depicted IMAC in **Table 1**) as described above.

For all purified enzymes, imidazole from the last IMAC step was removed by ultrafiltration through a membrane of 10-kDa cutoff (Amicon Ultra Centrifugal Filter Device; Millipore; Billerica, MA, USA). The purified and concentrated enzymes were washed three times with 10 mL 50 mM potassium phosphate buffer pH 7.0. As reported previously [17,18], *Am*PDH is mainly isolated in its reduced form. To fully oxidize *Am*PDH and its variants, the enzymes were re-oxidized with 5 mM 2,6-dichlorophenol-indophenol (DCIP). The solution was then passed through a PD-10 Desalting Columns (GE Healthcare) equilibrated with 50 mM potassium phosphate buffer pH 7.0 to remove DCIP. Then, the solution was sterile-filtered by passing through a membrane with a 0.22- μ m cutoff (Merck Millipore, Darmstadt, Germany) and finally diluted in sterile potassium phosphate buffer pH 7.0 to a protein concentration of approximately 20–50 mg mL⁻¹. Aliquots of 100 μ L were shock-frozen in liquid nitrogen and stored at -80°C. All variants were stored at 4°C after sterile-filtration.

5.2.5 Enzyme activity assay and molecular properties

PDH activity was determined by following the reduction of the ferrocenium ion at 300 nm and 30°C for 3 min as described previously [32] with the following adaptations: the standard reaction mixture contained 50 μ mol sodium phosphate buffer pH 7.5, 25 μ mol GLC, and 0.2 μ mol of ferrocenium hexafluorophosphate. Protein concentrations were determined according to Bradford with a pre-fabricated assay (BioRad, Hercules, CA, USA). SDS-PAGE and enzymatic deglycosylation with PNGase F were conducted as described previously [2,18].

Determination of molar absorption coefficients for *Am*PDH, H103A, and H556A were conducted with an adapted protocol as reported previously [24]. In brief, the enzyme solution was diluted in 50 mM potassium phosphate buffer pH 7.0 to an absorbance at 450 nm of about 2.5. To 100 μ L of that solution, either 900 μ L of the same buffer (for spectra of intact enzyme) or 900 μ L of 6 M

guanidine HCl (for spectra of denatured enzyme) were added. Concentrations of free FAD released from the denatured enzymes were determined in triplicates based on the free FAD molar absorption coefficient ($\epsilon_{450} = 1.1 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$). The extinction coefficient of FAD bound to following enzyme subunits were: $\epsilon_{463} = 9.5 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ for *Am*PDH, $\epsilon_{460} = 1.2 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for H103A, and $\epsilon_{465} = 1.1 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ for H556A. The low ϵ_{463} for *Am*PDH suggests a deprotonated flavin N3-atom [25,33].

5.2.6 Electronic circular dichroism spectroscopy

To monitor the overall fold and secondary structure elements and compare them to *Am*PDH, far UV Electronic circular dichroism (ECD) spectra were recorded as described previously [19]. ECD spectra of *Am*PDH and variants were recorded from 180 to 260 nm on a Chirascan CD Spectrophotometer (Applied Photophysics, Leatherhead, UK) which was flushed with nitrogen. The cuvette pathlength was 1 mm, the spectral bandwidth was set to 3 nm and the scan time per point to 10 s. The protein concentration was adjusted to approximately 4 μM with 50 mM potassium phosphate buffer pH 7.0.

5.2.7 ThermoFAD

*Thermo*FAD measurements were conducted to elucidate the thermal unfolding temperature T_m , as reported [23] utilizing the increased intrinsic fluorescence of the FAD cofactor upon thermal protein denaturation. Two replicas were measured, each containing 25 μL with 65 μM *Am*PDH or the respective variant in 40 mM Britton-Robinson buffer (pH 2–9, in steps of 1 pH units). Heating of samples was conducted from 20–95°C in steps of 0.5°C s^{-1} with an iCycler Thermal Cycler equipped with an MyiQ Real-Time PCR Optical Module (BioRad, Hercules, CA, USA), and the corresponding fluorescence was recorded. T_m was determined from the maximum of the first derivative of the obtained sigmoidal curve [23,34].

5.2.8 TCA/acetone precipitation

To determine whether FAD was covalently attached to the polypeptide, a solution of oxidized enzyme with an absorbance (450 nm) between 0.1–0.3 was prepared in 50 mM potassium phosphate buffer pH 7.0, and the oxidized spectrum was subsequently recorded from 250 to 650 nm with a U-3000 spectrophotometer (Hitachi, Tokyo, Japan). Precipitation was conducted as reported previously [18], by mixing the double-concentrated protein solution with 10% (v/v) TCA and 40% (v/v) acetone and incubation for 10 min on ice. After centrifugation at 13,000 rpm and 4°C for 5 min, the spectrum of the supernatant was recorded to identify the presence on noncovalently bound FAD.

5.2.9 Steady-state kinetics

Apparent kinetic constants for the electron donor GLC were measured utilizing the standard Fc^+ assay. Data for Fc^+ was collected in 100 mM borate buffer pH 8.5 and the concentration of the electron donor GLC was adjusted according to the K_m value of the corresponding variant: 25 mM (*nsAmPDH*, *recAmPDH*, and *AmPDH*), 50 mM (H103A, Q392A, and V511W), 250 mM (Y510A), and 500 mM (H556A, H556N, and V511F). The obtained data were fitted to the Michaelis-Menten equation by non-linear least squares regression with Sigma Plot 11 (Systat Software, Chicago, IL, USA), from which the kinetic constants were derived.

5.2.10 Molecular dynamics simulations and free energy calculations

The structure preparations were essentially performed as reported previously [15]. In short, two different binding poses (pose A and B) of GLC were considered, differing in a rotation of 180° about the axis defined by a line running through a point midway between the GLC atoms C5 and O5, and a point midway between atoms C2 and C3, allowing for GLC oxidation at C2 or C3. Only two protonation states of the active site histidines were considered, which had the most favorable binding energies and highest ligand stabilities of GLC in *AmPDH* in a previously reported comparison of the protonation states [15]. For pose A (as shown in **Figure 1**), His-512 and His-556 were both fully protonated (protonation state PP); for pose B, His-512 was fully protonated and His-556 was in its neutral state with a proton at Ne (protonation state PN). For *in silico* mutations to more bulky side chains (V511F and V511W), rotamers most similar to the wild-type were selected. During the 10 ns simulations, the side chains of those variants could rotate and adopt new conformations. For *in silico* mutations to less bulky residues, the atom positions of the side chains were derived from the wild-type.

The simulation setup was performed similarly as reported earlier [15]. The MD simulations were conducted with the GROMOS 11 software package [35] employing the 53A6 force field [36] in explicit solvent within a rectangular, periodic, and pre-equilibrated box of SPC water [37]. After energy minimization and equilibration [38], two independent 10 ns production runs (md1 and md2) were performed at constant pressure (1 atm) and temperature (300 K) by using the weak-coupling scheme [39] with coupling times of 0.5 and 0.1 ps, respectively. The isothermal compressibility was set to $4.575 \times 10^{-4} \text{ kJ}^{-1} \text{ mol nm}^3$, and two separate temperature baths were used for solute and solvent. The SHAKE algorithm was applied to constrain bond lengths [40] for solute and solvent allowing for 2-fs time-steps. Nonbonded interactions were calculated using a triple range scheme. Interactions within a short-range cutoff of 0.8 nm were calculated at every time step from a pair list that was updated every fifth step. At these points, interactions between 0.8 and 1.4 nm were also calculated explicitly and kept constant between updates. A reaction field [41] contribution was added to the electrostatic interactions and forces to account for a homogenous medium outside

the long-range cutoff using a relative dielectric constant of 61 as appropriate for the SPC water model [42]. Coordinate and energy trajectories were stored every 0.5 ps for subsequent analysis.

Binding free energies (ΔG_{bind}) relative to *Am*PDH ($\Delta\Delta G_{\text{bind}}$) were estimated from experimental K_m values ($\Delta\Delta G_{\text{bind}}^{\text{exp}}$) according to

$$\Delta\Delta G_{\text{bind}}^{\text{exp}} = k_B T \ln \left[K_m^{\text{variant}} / K_m^{\text{AmPDH}} \right] \quad \text{equation 1}$$

where $k_B T$ is the Boltzmann constant multiplied with the absolute temperature, or calculated from MD simulations ($\Delta\Delta G_{\text{bind}}^{\text{sim}}$) by employing the linear interaction energy (LIE) method [43] according to the following equation [44]

$$\Delta G_{\text{bind}}^{\text{sim}} = \beta \left(\langle V_{\text{lig-sur}}^{\text{EL}} \rangle_{\text{protein}} - \langle V_{\text{lig-sur}}^{\text{EL}} \rangle_{\text{free}} \right) + \alpha \left(\langle V_{\text{lig-sur}}^{\text{VdW}} \rangle_{\text{protein}} - \langle V_{\text{lig-sur}}^{\text{VdW}} \rangle_{\text{free}} \right) \quad \text{equation 2}$$

where angular brackets indicate ensemble averages, calculated over a simulation of the substrate bound to the protein (protein) or free in solution (free). $V_{\text{lig-sur}}^{\text{EL}}$ and $V_{\text{lig-sur}}^{\text{VdW}}$ represent the ligand-surrounding electrostatic and van der Waals interactions, respectively. α and β are parameters of the LIE equation. $\Delta\Delta G_{\text{bind}}^{\text{sim}}$ was obtained by subtracting $\Delta G_{\text{bind}}^{\text{sim}}$ of wild-type *Am*PDH from $\Delta G_{\text{bind}}^{\text{sim}}$ of the respective variant. The different poses (pose A and B) and runs (md1 and md2) were averaged as described in [45].

5.2.11 GC-MS measurements of GLC reaction products

Stock solutions of glucose, 2-keto-GLC and meso-erythritol were prepared by dissolving appropriate amounts of solid standard in LC-MS grade water in amber LC vials and kept at -40°C . Working solutions were prepared daily by diluting the stock solution in LC-MS grade water. Ethoxymation solution was prepared daily by dissolving 18.7 mg of ethoxylamine hydrochloride in 1 mL water-free pyridine. A nonpolar HP-1MS column (60 m length \times 250 μm i.d. \times 0.25 μm thickness film, 100 % dimethylpolysiloxane), and a guard column (5 m length \times 320 μm i.d. without stationary phase) from Agilent Technologies (Santa Clara, CA, USA) were used for separation. EZ-2 Envi SpeedVAC from Genevac Inc (Stone Ridge, NY, USA) was used for drying the standards and samples. Gas chromatography–mass spectrometry (GC-MS) was carried out with an Agilent Technologies GC-EI-MS machine consisting of an Agilent Technologies 7820A gas chromatograph and an Agilent Technologies 5975 mass selection detector. Gas chromatography in combination with chemical ionization time of flight mass spectrometry (GC-CI-ToFMS) was conducted on an Agilent Technologies 7200 GC-MS. Sample preparation and derivatization was adopted from [46].

The derivatization product solution was kept at 4°C in an auto-sampler, of which 1 μL was injected into a nonpolar HP-1MS column with an Agilent Technologies guard column with pulsed injection mode at 30 psi. The temperature of the injector and the transfer line were kept constant at 250°C

and 280°C, respectively. Helium gas (99.9999% purity grade) was used as carrier gas with a flow rate of 1.2 mL min⁻¹ in constant flow mode. The column oven was kept at 70°C for 1 min and heated up to 200°C (20°C min⁻¹), then increased to 240°C (2.5°C min⁻¹), and finally to 310°C (20°C min⁻¹) where it was kept for 1 min to re-condition the column. For GC-CI-QToFMS, temperature, pressure of methane gas, emission current, electron energy as well as QToFMS parameters were optimized as reported previously [47]. The mass range in both cases was scanned from 70 to 700 amu. An n-alkane (C8-C40) retention index in ethoxylamine and MSTFA solution was injected into GC-EI-MS in order to calculate the modified retention indices according to [48]. Cross-check mass fragmentation, pathway of electron ionization mass spectra, and chemical ionization was carried out with Mass Frontier version 7.0 SP1 (Thermo Scientific, USA). Quantification of 2-keto-GLC was carried out with external standards. For 3-keto-GLC and 2,3-diketo-GLC no standards were commercially available. Therefore, semiquantitative detection was performed by comparing the peak areas of the corresponding derivatization-products with 2-keto-GLC at the same mass to charge ratio.

5.2.11.1 Measurement of GLC reaction products *via* GC-EI-MS and GC-CI QToFMS

A total ion chromatogram of a sample analyzed by GC-EI-MS is presented in the **Figure S1A**. The standard electron ionization (70 eV) mass spectra of the two-step derivatization product of 2-keto-GLC was confirmed *via* standard EI mass spectra with low resolution, which gave an m/z of 552.4 (M+, <1%), 537.4 ([M-CH₃]+, 5%), 417.2 (5%), 347.2 (18%), 307.2 (40%), 277.1 (12%), 246.1 (5%), 217.1 (60%), 147.1 (35%), 103.1 (65%), and 73.1 (100%, trimethylsilylium ion). Mass spectra of the two-step derivatization product of 2-keto-GLC was successfully cross-checked with published data [49]. In addition, the fragmentation pathway of the two-step derivatization products of 2-keto-GLC and 3-keto-GLC were predicted with Mass Frontier version 7.0SP1 from Thermo Scientific (CA, USA). For 3-keto-GLC, the mass spectrum of the two-step derivatization products was similar to 2-keto-GLC, except for one fragment derived from 3-keto-GLC with an m/z of 174.1, which differed significantly between two isomers. With this fragment and the protonated ion in methane, chemical ionization spectra were used to screen for this compound in the enzyme reactions. However, a high degree of fragmentation during electron ionization and unspecific fragments in standard electron ionization were observed, preventing the distinction between 2-keto-GLC and 3-keto-GLC by only EI or CI mass spectra. Because of the lack of commercially available standards for 3-keto-GLC and 2,3-diketo-GLC, the discrimination between the different keto-GLC derivatives was only possible by combining chemical ionization mass spectra and electron ionization mass spectra with the fragmentation pathway. The molecular formula generation (MFG) was calculated *via* the compound identification function in the MassHunter software. Elements (C, N, O, and Si) and number of atoms were limited in accordance with derivatization products of mono- and diketo-GLC.

The chemical formula of the derivatization product of 2-keto-GLC and 3-keto-GLC is $C_{22}H_{52}O_6N_2Si_4$ with an accurate mass of 552.2897 Da, and $C_{21}H_{47}O_6N_3Si_3$ for 2,3-diketo-GLC with an accurate mass of 521.2767 Da. The m/z of the compounds was increased by 1.0080 Da after ionization by methane as chemical reagent gas. The theoretical isotopologue fractions of derivatization product ions of mono- and diketo-GLC were calculated *via* enviPat (<http://www.envipat.eawag.ch/index.php>). High-resolution spectra of ethoxymation-trimethylsilylation of mono- and diketo-GLC employing methane gas ionization are shown in **Figure S1B**. For the investigated mono- and diketo-GLC, the mass accuracy of the quasi-molecular ion was below 10 ppm and the experimental isotopologue fractions were in agreement with the theoretical mass (**Table S2**). Identification of 2,3-diketo-GLC was accomplished by its retention time during chromatography. In addition, the molecular formula generation function in qualitative MassHunter software with atoms filter C, H, N, O and Si was set to 19-23, 47-55, 2-4, 5-8, and 2-5, respectively. The comparison of selected compounds with the molecular formula generation was carried out not only by m/z and isotopologue fraction but also according to the space between the isotopic pattern. The match scores were 93.9 and 96.4 for mass and isotopologue fraction distribution, respectively. **Figure S1C** shows the theoretical and experimental isotopologue fractions for 2,3-diketo-GLC.

5.2.12 Rapid reaction experiments

Reactions were carried out similarly as described previously [50]. In short, 50 mM potassium phosphate buffer pH 7.0 at 4°C, a TgK Scientific model SF-61DX or a TgK Scientific model SHU-61SX2 (TgK Scientific, Bradford-on-Avon, UK) stopped-flow spectrophotometer in single-mixing mode and a dead-time of 2 ms were used. The stopped-flow instrument was made anaerobic by flushing the flow system with an anaerobic buffer solution containing 0.5 mg/mL dithionite in 50 mM potassium phosphate buffer pH 7.0 and equilibrated in that solution over night. Before starting the experiments, the flow system of the instrument was washed three times with anaerobic 50 mM potassium phosphate buffer pH 7.0.

The enzymes obtained from purification contained partially reduced populations. To prepare fully oxidized enzymes, 5 mM 2,6-dichlorophenol-indophenol (DCIP) was used as described at the end of section 5.2.4. To prepare a reduced enzyme solution, an anaerobic oxidized enzyme solution was stoichiometrically reduced in an anaerobic glove box with a 300 μ M GLC solution in potassium phosphate buffer pH 7.0, which was monitored spectrophotometrically. All substrate concentrations used were more than 5-fold excess of the enzyme concentration (30 μ M) to ensure pseudo-first order conditions.

To study the kinetic reduction of wild type *Am*PDH by sugar substrates, the reaction was performed aerobically, because the enzyme has previously been shown to be reoxidized very slowly with oxygen [17]. Anaerobic conditions were used for stopped-flow experiments of variant H556A,

which was shown to have a 3.7-fold elevated oxygen reactivity compared to *Am*PDH (see section 5.3.7). Reoxidation experiments with 1,4-benzoquinone were also performed anaerobically to prevent side-reactions of 1,4-benzoquinone with oxygen.

Observed rate constants (k_{obs}) were calculated from exponential fits to kinetic traces using the software Kinetic Studio (TgK Scientific, Bradford-on-Avon, UK) and Program A (Rong Chang, Jung-yen Chiu, Joel Dinverno, and D.P. Ballou, University of Michigan, MI, USA). Obtained k_{obs} values were plotted against the substrate concentrations and the apparent bimolecular rate constant (k_{app}) was calculated from the slope. Simulations were performed by numerical methods with Runge-Kutta algorithms implemented in the software Berkeley Madonna 8.3 (University of Berkley, CA, USA) with a time step of 2×10^{-4} s. The kinetic models utilized to simulate the reductive- and oxidative half-reaction are listed in the results section.

5.2.13 Redox potential determination

The redox potentials of *Am*PDH, H103A, and H556A were measured according to Massey's method [27] using xanthine and xanthine oxidase as the reduction system [7,51]. The enzymes were fully oxidized with 5 mM 2,6-dichlorophenol-indophenol (DCIP) as described at the end of section 5.2.4. Solutions of fully oxidized enzyme (*Am*PDH, H103A, or H556A), benzyl viologen, the standard dye, xanthine, and xanthine oxidase (side arm) were pipetted together to a final volume of 1 mL in a specially designed cuvette equipped with two side arms and a stopcock. Anaerobiosis was established by repeated cycles of evacuation and flushing with oxygen-free nitrogen. Afterwards, the reaction was started by adding xanthine oxidase from the side arm. The reduction was monitored by recording spectra with an Agilent Technologies 8453 diode array spectrophotometer (Santa Clara, CA, USA).

Phenazine methosulfate ($E_m^o = 80$ mV) was used as standard dye for *Am*PDH. Enzyme and dye were monitored at 463 nm (*Am*PDH) and 387 nm (phenazine methosulfate) during their slow reduction over a period of approximately 8 h. *Am*PDH has no isosbestic point at 387 nm but a significant absorbance, which changes with altering redox-states of the enzyme. Phenazine methosulfate has a small absorbance at 463 nm, which hardly changes during the reduction process of the dye. Consequently, the absorbance at 463 nm and 387 nm had to be analyzed with respect to a mixed population of enzyme and dye. For this, following equations were used

$$\epsilon_{463 \text{ nm}}^{\text{E}} c^{\text{E}} + \epsilon_{463 \text{ nm}}^{\text{D}} c^{\text{D}} = A_{463 \text{ nm}} \quad \text{equation 3}$$

$$\epsilon_{387 \text{ nm}}^{\text{E}} c^{\text{E}} + \epsilon_{387 \text{ nm}}^{\text{D}} c^{\text{D}} = A_{387 \text{ nm}} \quad \text{equation 4}$$

whereas superscript E abbreviates 'Enzyme' and D 'Dye' at the wavelength indicated in the subscript, ϵ represents the extinction coefficient, c the (unknown) concentration. This system of two equations was solved to yield the concentration of the enzyme or dye. These concentrations

were subsequently used to calculate the ratio of reduced and oxidized species at various stages during the reduction process. Afterwards, the midpoint potential (E_m^o) of *Am*PDH was determined using the standard Nernst equation as described elsewhere [52].

For variants H103A and H556A, methylene blue ($E_m^o = 11$ mV) was used as standard dye. The variants and dye were monitored at 460 nm (H103A), 465 nm (H556A), and 664 nm (methylene blue), whereas the monitored absorbance of variant H103A or H556A did not interfere with the absorbance of the dye and *vice versa*. Consequently, the concentrations of dye and variant could be readily obtained by the absorbance at the indicated wavelength. For both variants, these concentrations were subsequently used to calculate the ratio of reduced and oxidized species at various stages during the reduction process. Since for both variants an intermediate species with an absorbance maximum of 422 nm (H103A) or 368 nm (H556A) was observed, which points towards the red semiquinone, subsequent calculations employing the standard Nernst equation [52] yielded the semiquinone/reduced enzyme half potential ($E_2^{o'}$). $E_2^{o'}$ could be used to calculate E_m^o for both variants according to following equations [51,53,54]

$$E_1^{o'} - E_2^{o'} = 2 (2.303 RT/F) \log[2 M/(1 - M)] \quad \text{equation 5}$$

$$E_1^{o'} + E_2^{o'} = 2 E_m^o \quad \text{equation 6}$$

where M = the maximum fraction of thermodynamically stable semiquinone formed, $T = 298$ K, E_m^o = the two-electron reduction potential, $E_1^{o'}$ = the oxidized enzyme/semiquinone half potential, and $E_2^{o'}$ = the semiquinone/reduced enzyme half potential. The amount of thermodynamically stable semiquinone (M) was quantified in mixtures with similar compositions as described above. In these mixtures, the standard dye was replaced by 50 μ L of a 2 mM allopurinol solution, a potent xanthine oxidase inhibitor [55], which was pipetted into the second side arm of the cuvette. The xanthine and xanthine oxidase reduction system was stopped after semiquinone absorption reached its maximum (SQ_{MAX}) at 422 nm (H103A) or 368 nm (H556A) by adding the allopurinol solution from the side arm of the cuvette. The absorbance at 422 nm (H103A) or 368 nm (H556A) was recorded again after letting the reaction mixture stand at room temperature in the dark overnight (SQ_{ON}). Finally, M was obtained by calculating SQ_{ON}/SQ_{MAX} .

5.2.14 O₂ reactivity

The steady-state oxygen reactivity for all purified enzymes was determined in quadruplicates with a fluorimetric Amplex Red/horseradish peroxidase assay as reported previously [18], utilizing the calibration curve as reported therein. In the current study, 25 and 100 μ M of the electron donor GLC were used to compensate for rather high apparent K_m values for GLC for some variants. The obtained oxygen reactivities were essentially the same for both GLC concentrations. Therefore, only oxygen reactivities for 25 μ M GLC are reported here, which is in line with previously published

results [18]. The measured H_2O_2 generation (in $\text{min}^{-1} \text{mg}^{-1}$ enzyme) was converted *via* the calibration curve into the oxygen reactivity, which is given in $\mu\text{M O}_2 \text{min}^{-1} \text{mg}^{-1}$ enzyme.

5.3 Results and discussion

5.3.1 Cloning, expression, and purification

We chose six active site residues as targets for mutagenesis and structure / function studies based on the crystal structure of *Agaricus meleagris* PDH [17] (*Am*PDH; PDB code 4H7U) and on recent findings employing molecular dynamics (MD) simulations [15]: His-103, Gln-392, Tyr-510, Val-511, His-512, and His-556 (**Figure 1**). These recent studies showed that the selected residues are involved in substrate binding (Gln-392, Tyr-510, Val-511, His-556), catalysis (His-512, His-556), or in covalently tethering the FAD cofactor to the protein (His-103). Site-directed mutagenesis [56] (primer pairs are listed in **Table S1**) to introduce alanine [57] or other, rationally selected amino acids was utilized to create variants for functional characterizations. In a previous *in silico* study [15] we proposed that the backbone oxygen atom of Val-511 is important for substrate binding. This motivated us to generate variants V511F and V511W, which both have bulky side chains. Both active site histidines (His-512 and His-556) were previously suggested to act as catalytic bases [15,17,18], however, these two histidines have not been studied in any detail. To this end, we generated alanine variants H512A and H556A. The closely related enzyme pyranose 2-oxidase from the fungus *Trametes multicolor* (*Tm*P2O) carries an asparagine at position 593 [7,58], corresponding to the histidine at position 556 in *Am*PDH. Because this is the most striking difference in the active site architecture between *Tm*P2O and *Am*PDH, we generated the asparagine variant H556N for *Am*PDH as well. Other variants that were studied in an alanine-scanning approach include H103A, Q392A and Y510A.

All proteins were heterologously expressed using *P. pastoris*, which proved to be superior for the production of recombinant *Am*PDH to other microbial expression systems [3]. H103A, Q392A, Y510A, V511F, V511W, H512A, and H556N were produced in 5-L scale as reported previously [18], whereas wild-type recombinant *Am*PDH and H556A were obtained in 30-L scale in principle as described in [3]. Purification according to a novel purification scheme employing cross-flow filtration and immobilized metal affinity chromatography (IMAC) is described in detail in the materials and methods section, the corresponding purification data are presented in **Table 1**.

Table 1: Purification of recombinantly produced AmPDH and its variants. Abbreviations of purification steps: CE, crude extract, corresponding to the supernatant of the fermentation broth; UF, ultrafiltration in cross-flow mode with a membrane-cutoff of 10 kDa; IMAC, immobilized-metal affinity chromatography, where IMAC I depicts the capture step with a 65-mL column, and IMAC or IMAC II the polishing step with 5-mL columns; HIC, hydrophobic interaction chromatography with a 200-mL column.

Variant	Purification step	Volume [mL]	Protein [mg]	Activity [U]	Specific activity [U/mg]	Recovery [%]	Purification [-fold]
AmPDH^a	CE	20,000	6,760	52,800	7.80	100	1.00
	UF	1,850	7,380	45,200	6.12	85.7	0.78
	IMAC I	18.0	819	30,300	37.0	57.4	4.74
	HIC	190	402	24,100	60.0	45.8	7.69
	IMAC II	5.00	360	16,800	46.7	31.9	5.99
H103A^b	CE	3,220	1,840	1,970	1.07	100	1.00
	HIC	310	431	2,260	5.24	115	4.90
	IMAC	1.50	174	1,830	10.6	93.3	9.88
Q392A^b	CE	3,040	1,870	4,340	2.32	100	1.00
	HIC	510	525	3,600	6.86	82.9	2.95
	IMAC	1.08	93.1	1,870	20.1	43.1	8.65
Y510A^b	CE	3,080	1,670	575	0.34	100	1.00
	HIC	350	344	319	0.93	55.5	2.70
	IMAC	0.94	35.8	211	5.88	36.7	17.1
V511F^b	CE	2,800	1,090	336	0.31	100	1.00
	HIC	565	435	186	0.43	55.5	1.38
	IMAC	0.50	7.78	31.8	4.08	9.45	13.2
V511W^b	CE	3,500	1,640	2,100	1.28	100	1.00
	HIC	500	441	2,260	5.13	108	4.00
	IMAC	0.82	52.8	710	13.5	33.9	10.6
H512A^b	CE	3,500	1,620	165	0.10	100	1.00
	HIC	214	361	4.45	0.01	2.71	0.12
	IMAC	1.00	20.8	0.02	0.00	0.01	0.01
H556A^a	CE	22,000	5,040	2,490	0.49	100	1.00
	UF	1,300	4,310	1,550	0.36	62.3	0.73
	IMAC I	100	483	1,370	2.83	54.8	5.72
	HIC	270	441	1,320	3.00	53.1	6.07
	IMAC II	7.50	410	1,070	2.60	42.8	5.25
H556N^b	CE	3,880	1,390	206	0.15	100	1.00
	HIC	410	240	80.5	0.33	39.1	2.27
	IMAC	1.00	32.5	55.6	1.71	27.0	11.6

^a Produced in *P. pastoris* in 30 L scale as reported previously [3]

^b Produced in *P. pastoris* in 5 L scale as reported previously [18]

5.3.2 Molecular properties

After chromatographic purification, fractions of highest purity were combined for each variant, resulting in a final pool with an apparent homogeneity of >98% as judged by SDS-PAGE (**Figure 2**). *Am*PDH and all variants showed a broad smear around 90 kDa on SDS-PAGE. Deglycosylation with PNGase F under denaturing conditions resulted in one distinct band of about 64 kDa, indicating a degree of glycosylation of roughly 30%, which is in good agreement with previously published work [3,18]. This rather high degree of glycosylation is caused by glycan additions of the high-mannose type by *P. pastoris* [59], as *Am*PDH purified from its native source *Agaricus meleagris* has a sugar content of about 7% [2].

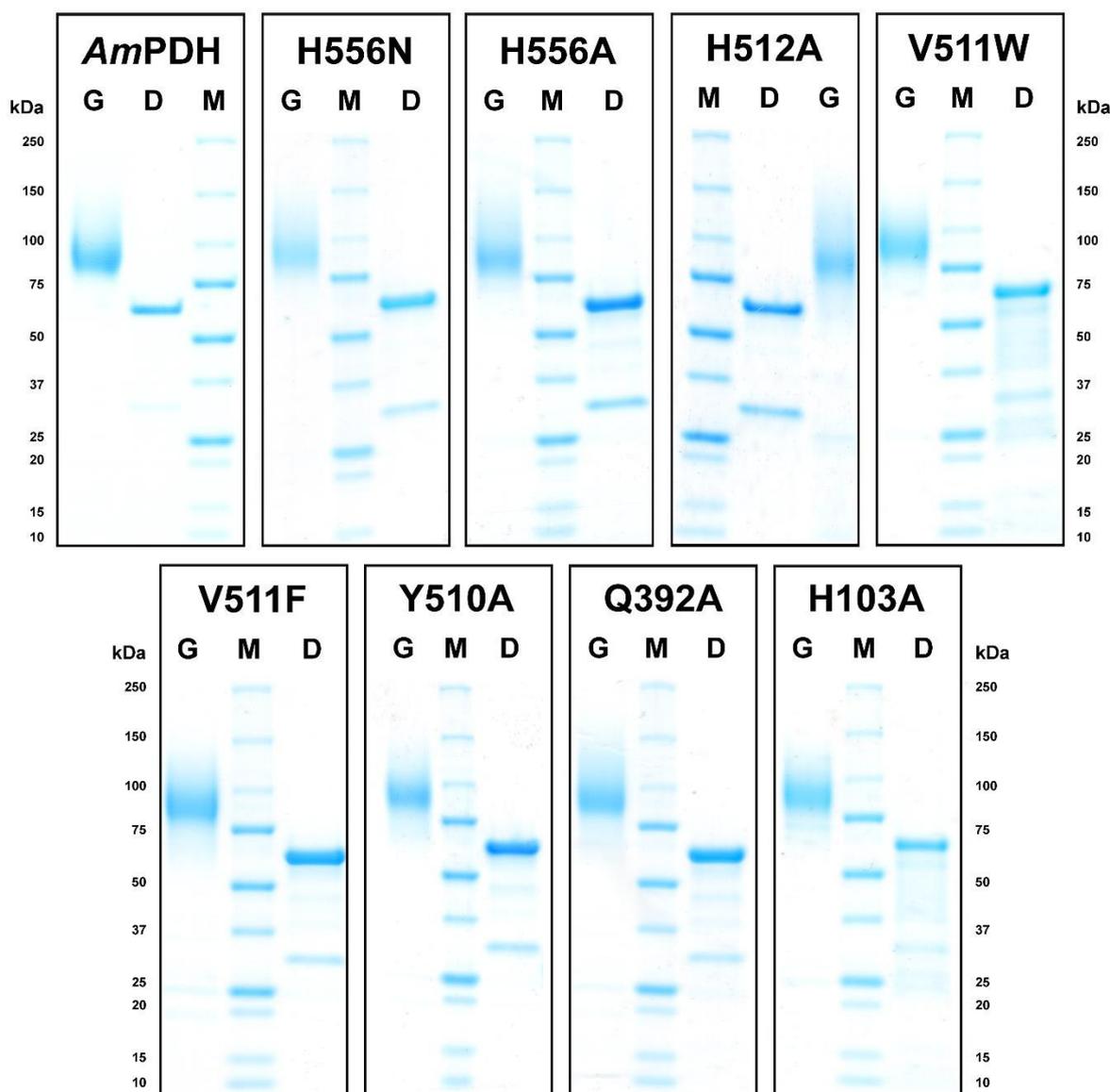


Figure 2: SDS-PAGE of *Am*PDH and its variants expressed in *P. pastoris* as reported previously [3,18] and purified by a novel purification scheme employing immobilized metal affinity chromatography (IMAC). Figure legend: M – molecular marker (Precision Plus Protein Standard, BioRad); G – glycosylated enzyme; D – deglycosylated enzyme (with PNGase F). 3–6 μ g protein was loaded onto each lane. The band around 36 kDa in lanes depicted ‘D’ is derived from PNGase F, which was used for deglycosylation.

To assess the overall fold of the variants, far-UV ECD spectra were recorded and compared to wild-type recombinant *Am*PDH (**Figure 3**). All proteins show very similar far-UV ECD spectra, indicating that they have very similar secondary structure [22]. Consequently, the mutations did not affect proper protein folding. Distinct minima at 208 and 222 nm indicate a dominating α -helical content [22,60], which agrees nicely with secondary structure elements found in the PDH crystal structure [17].

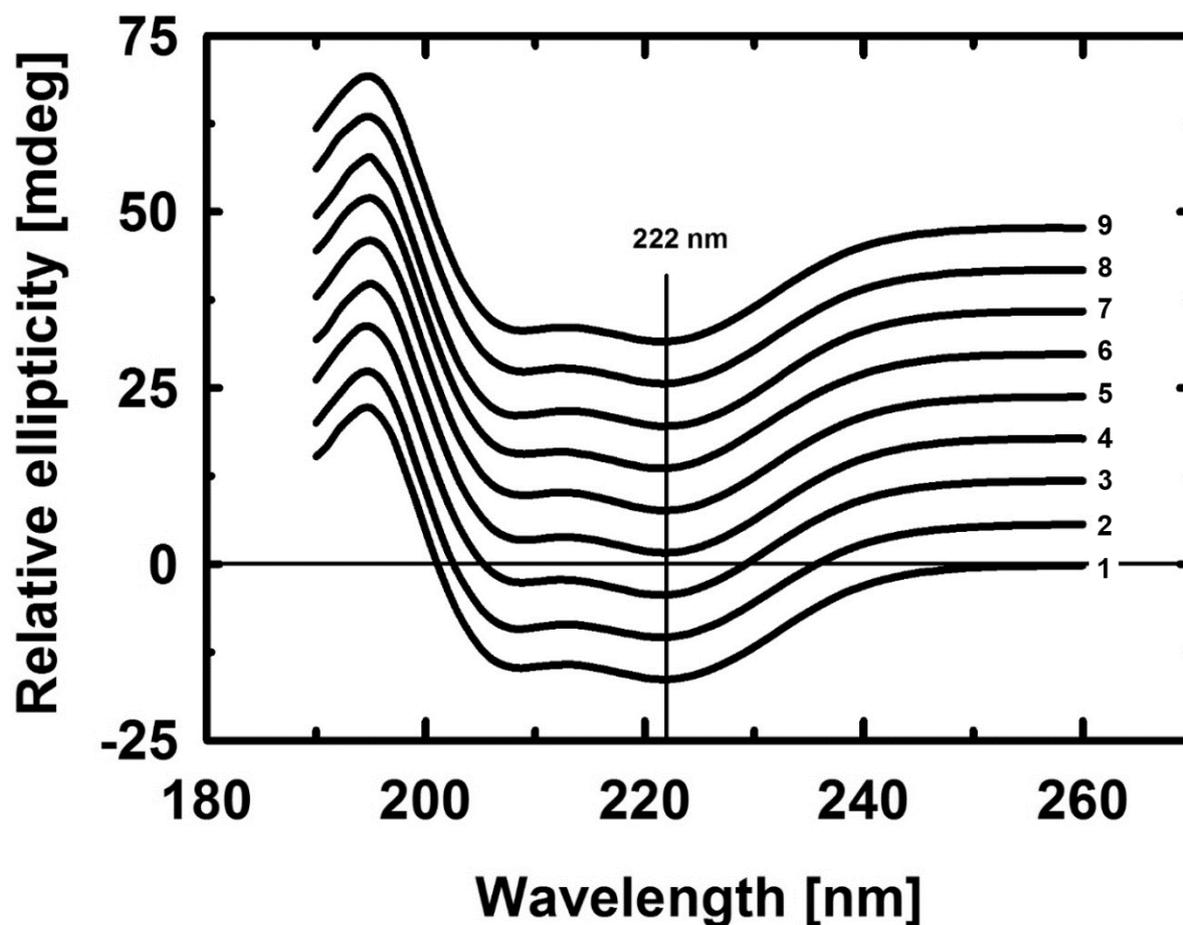


Figure 3: Far-UV ECD spectra of 4 μ M protein solutions in 50 mM potassium phosphate buffer at pH 7.0. For baseline correction, the ECD spectrum of 50 mM potassium phosphate buffer at pH 7.0 was subtracted from all protein ECD spectra. Numbering scheme for assigning *Am*PDH variants to the spectra, which were multiplied with the numbers in brackets to normalize the spectra to the 222 nm value of H103A: 1 – *Am*PDH (0.947); 2 – H103A (1); 3 – Q392A (0.936); 4 – Y510A (0.979); 5 – V511F (1.066); 6 – V511W (1.226); 7 – H512A (0.845); 8 – H556A (1.017); 9 – H556N (0.949). To arrange the ECD spectra successively, the value 6 was added to spectrum 2 (H103A), 12 was added to spectrum 3 (Q392A), etc.

To investigate the effect of the introduced mutations on thermostability, the unfolding temperature (T_m) was determined with the *ThermoFAD* method [23] in the range of pH 2 to 9 (**Figure 4**). As indicated by T_m , all variants show their highest thermostability in the range of pH 5–6. A T_m value of 73.5°C was determined for recombinant wild-type *Am*PDH, which is in good agreement with previously published *ThermoFAD* data [19]. Interestingly, variant H103A has a T_m of 69.5°C, which is 5.7°C higher compared to the H103Y variant studied previously [19]. Apparently, the introduction of the bulkier tyrosine side chain at this position destabilizes the protein more significantly. The biggest decrease in T_m for all studied variants was observed for H512A ($T_m = 67.0^\circ\text{C}$), which is 6.5°C lower than the value of *Am*PDH, indicating an importance for His-512 in stabilizing the active site architecture. T_m values ranging from 68.8 to 71.8°C were determined for the other variants, indicating that the amino acid replacements chosen have no major effect of thermostability and do not destabilize any of the variants to an extent that makes further studies complex.

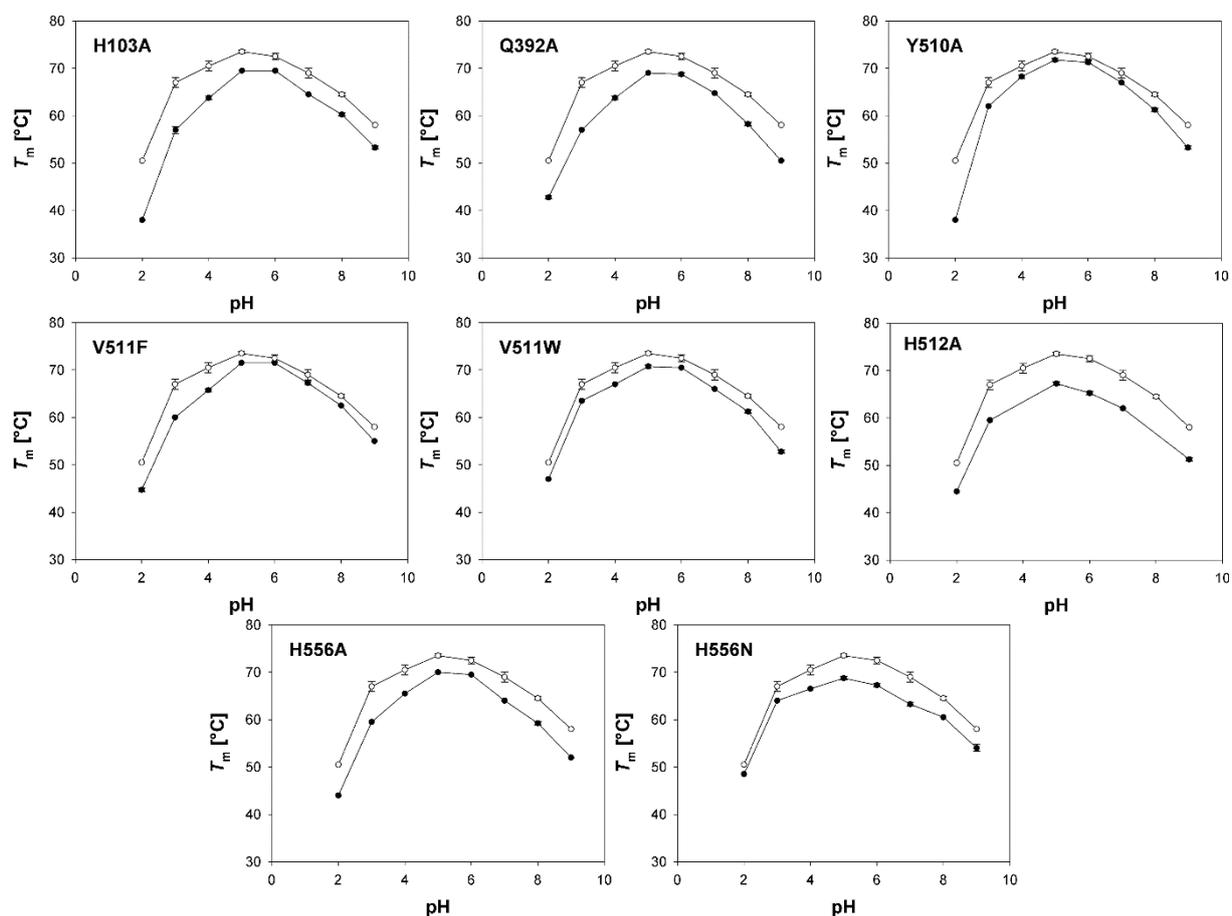


Figure 4: T_m values obtained from *ThermoFAD* experiments of 65 μM *Am*PDH and its variants in 40 mM Britton-Robinson buffer (pH 2–9). The upper line (white circles) represents the T_m value of wild-type recombinant *Am*PDH, and the lower line (black circles) the variant indicated. Because of limited H512A amounts available, data could only be recorded for pH 2, 3, 5, 6, 7, and 9. Data shown are the mean value of at least two independent experiments \pm SD.

Studies on the closely related enzyme *TmP2O* showed that active site variants can contain mixed populations of (non-)covalently bound FAD [24,61]. This is also true for amino acid replacements not involved in the covalent linkage as a result from a decrease of the overall positive charge around the flavin N1 region. Consequently, covalent FAD attachment was tested for all generated variants using protein precipitation with 10% trichloroacetic acid (TCA) and 40% acetone. Spectra of the protein solution before and after precipitation were recorded and are presented in **Figure 5**.

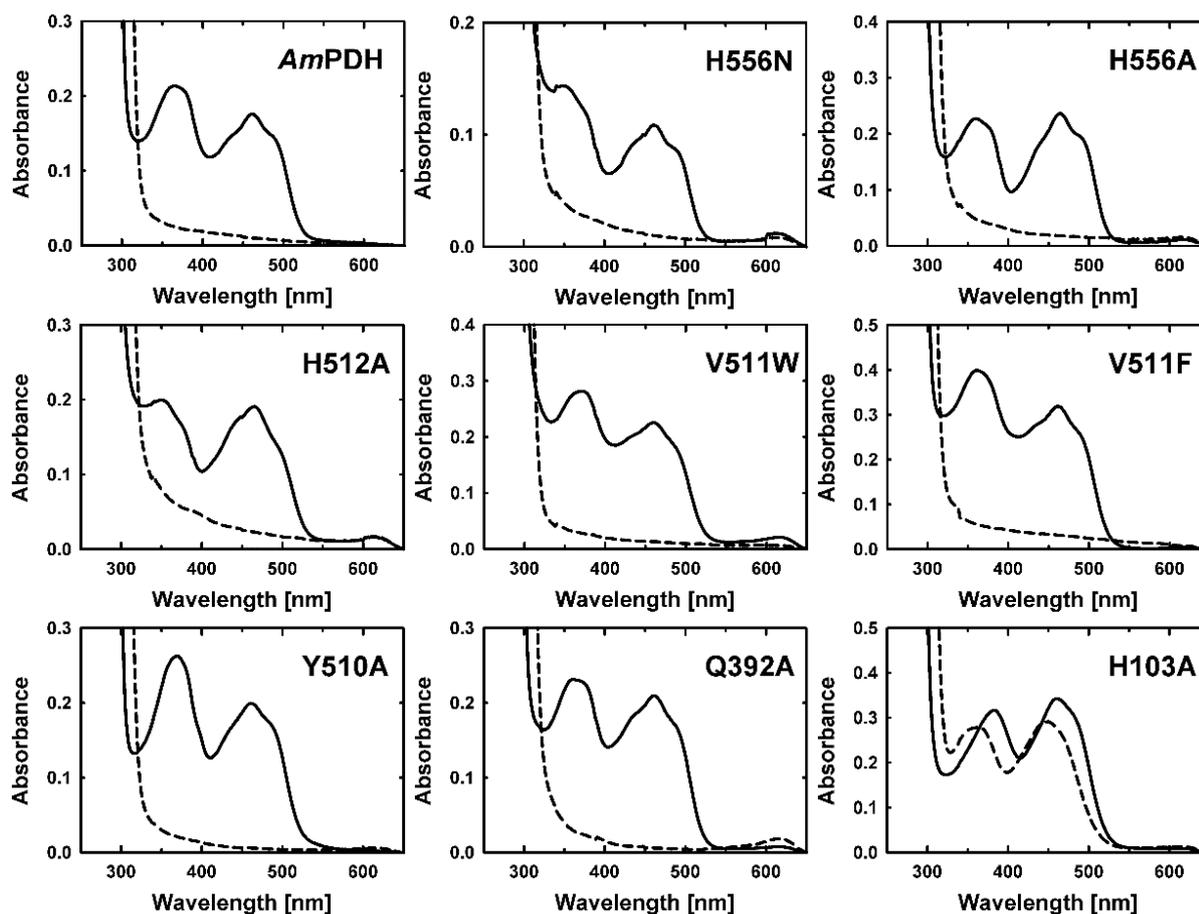


Figure 5: TCA/acetone precipitation of AmPDH and its variants as indicated in the upper right corner of each plot. Initially, UV/VIS spectra of fully oxidized enzymes were taken (solid line). After adding 10% (v/v) TCA and 40% (v/v) acetone to the double-concentrated solution and incubation for 10 min on ice, UV/VIS spectra of the supernatant were recorded again (dashed lines).

For variant H103A, lacking the covalent linkage to FAD, 89% of the protein released the FAD cofactor into the supernatant after precipitation. This percentage nicely agrees with variant H167A in *TmP2O* (also lacking the covalent linkage to FAD), for which 91% of the FAD was detected in the supernatant after precipitation [61]. For AmPDH and all other variants, <4% FAD was found in the supernatant, indicating covalently bound FAD cofactor. The UV-VIS spectra of the oxidized, non-precipitated enzyme solutions (black lines in **Figure 5**) show a hypsochromic (blue)

shift of the flavin peak at about 370 nm for all variants with a covalently attached FAD. This shift was observed previously for *Am*PDH [18] and indicates the 8 α -substituted flavin [25,62], which in the case of PDH is an 8 α -*N*³-histidyl-FAD [17,26]. Free FAD has a characteristic absorption peak at 450 nm [25], which is red-shifted in *Am*PDH and all variants. This bathochromic shift of the 450 nm peak upon FAD integration into a protein has been described earlier *e.g.* for monomeric sarcosine oxidase [63], which is caused by more extended delocalization of the π -electron system of the aromatic ring system in FAD [64]. For H103A, the maximum of this absorption peak blue-shifts from 460 to 450 nm after protein precipitation (from solid to dashed line in **Figure 5**), as expected for free FAD.

The results presented in this section clearly demonstrate that the produced enzymes are of high purity, overall possess an identical fold as *Am*PDH, are stable, and show UV-VIS spectra characteristic for flavoproteins.

5.3.3 Apparent steady-state kinetics: experiments rationalized by MD simulations and free energy calculations

To probe whether an amino acid is involved in substrate binding or turnover, apparent steady-state constants were determined for the electron donor GLC and the electron acceptor ferrocenium (Fc^+) (**Table 2**). For Fc^+ , variants can be classified in two groups: those with K_m values similar to the wild-type (H103A, V511W, H556A), and those with a K_m value that is at least twice as high (Q392A, Y510A, V511F, H556N). For all variants, the k_{cat} values are between 20-40 s^{-1} and thus significantly lower than for the wild-type, with the exception of Q392A, which shows an elevated k_{cat} of 111 s^{-1} . Since the reductive half-reaction was described to be the main determinant for the catalytic efficiency of *Am*PDH [19], and because of the unknown Fc^+ binding site as well as electron transfer path between the reduced enzyme and the electron acceptor, one should be careful in over-interpreting the apparent steady-state kinetic data obtained for Fc^+ , and we therefore refrain from drawing any further mechanistic conclusions.

In a next step, the apparent steady-state kinetics for GLC was analyzed in detail. Studies employing molecular dynamics simulations [15] and semi-rational protein engineering [18] suggested that the active site histidines His-512 and His-556 can act both as catalytic base during GLC oxidation. H512A shows roughly a 10-fold increase in K_m ($K_{m,\text{GLC}}$ of 6.19 mM) and a 41,200-fold lower k_{cat} ($k_{\text{cat},\text{GLC}}$ of 0.001 s^{-1}) compared to recombinant wild-type *Am*PDH, indicating its functional importance. These values are in perfect agreement with studies on the H548N variant of closely related *Tm*P2O, where His-548 corresponds to His-512 in PDH: for H548N, $K_{m,\text{GLC}}$ from 0.72 to 1.10 mM, whereas $k_{\text{cat},\text{GLC}}$ was dramatically reduced 46,000-fold compared to wild-type *Tm*P2O [7]. In more detailed studies, His-548 in *Tm*P2O was identified as the general base abstracting the 2-hydroxyl proton of GLC and initiating the hydride transfer from the substrate to the flavin [24]. For *Am*PDH variants involving His-556, $k_{\text{cat},\text{GLC}}$ is diminished by a factor of 3 (H556A) or 7

(H556N) when compared to *Am*PDH, indicating that the substrate is still efficiently oxidized. The apparent $K_{m, \text{GLC}}$ values for both, H556A and H556N, are elevated by two orders of magnitude, indicating that GLC binding is severely hampered. These findings unequivocally prove for the first time, that only His-512 takes on the role as a general base in *Am*PDH abstracting a proton from a sugar substrate, whereas His-556 seems to be mainly involved in substrate binding. Consequently, the substrate promiscuity of *Am*PDH cannot be explained by the presence of two active site histidines that can both act as a general base, but must be attributed to other factors, such as a diverse H-bonding pattern and the dominance of relatively unspecific interactions (*e.g.* van der Waals) for substrate binding in different poses as observed before [15]. This is in agreement with previous studies on the GMC family members aryl-alcohol oxidase (AAO, PDB: 3FIM) where active site histidines His-502 (corresponding to His-512 in PDH) and His-546 (corresponding to His-556 in PDH) were shown to have identical roles as proposed for *Am*PDH, namely proton abstraction and substrate interaction, respectively [4]. For glucose 1-oxidase (GOX, PDB: 1GAL), similar observations were obtained by experimental and computational studies for active site histidines His-516 (corresponding to His-512 in PDH) and His-559 (corresponding to His-556 in PDH) [65].

Table 2: Apparent steady-state kinetic constants of *Am*PDH and its variants for the electron donor D-glucose (GLC) and the electron acceptor ferrocenium (Fc^+). GLC was probed utilizing the standard Fc^+ assay (50 mM sodium phosphate buffer pH 7.5, 30°C, fixed value of 0.2 mM Fc^+ as electron acceptor). Fc^+ was tested in 100 mM borate buffer pH 8.5, 30°C. The fixed GLC concentration was adjusted according to the K_m value of the corresponding variant: 25 mM (*nsAm*PDH, *recAm*PDH, *Am*PDH), 50 mM (H103A, Q392A, V511W), 250 mM (Y510A), or 500 mM (H556A, H556N, V511F).

Variant	GLC			Fc^+		
	K_m [mM]	k_{cat} [s^{-1}]	k_{cat} / K_m [$\text{mM}^{-1} \text{s}^{-1}$]	K_m [mM]	k_{cat} [s^{-1}]	k_{cat} / K_m [$\text{mM}^{-1} \text{s}^{-1}$]
ns<i>Am</i>PDH^a	0.82 ± 0.03	45.9 ± 0.3	57.5	0.13 ± 0.03	104 ± 8	802
rec<i>Am</i>PDH^b	0.69 ± 0.09	37.8 ± 1.1	54.8	0.16 ± 0.04	130 ± 11	812
<i>Am</i>PDH^c	0.82 ± 0.01	42.2 ± 0.2	51.5	0.17 ± 0.02	135 ± 7	776
H556N	97.5 ± 4.1	6.40 ± 0.08	0.07	0.47 ± 0.04	22.5 ± 1.3	47.8
H556A	90.0 ± 3.5	12.8 ± 0.2	0.14	0.17 ± 0.01	20.4 ± 0.5	124
H512A	6.19 ± 1.24	0.0010 ± 0.0001	0.0002	ND ^d	ND ^d	ND ^d
V511W	3.73 ± 0.17	18.6 ± 0.2	5.00	0.11 ± 0.01	38.0 ± 1.4	352
V511F	106 ± 5	19.9 ± 0.3	0.19	0.39 ± 0.03	26.4 ± 1.4	66.8
Y510A	60.7 ± 2.0	16.0 ± 0.2	0.26	0.49 ± 0.03	28.8 ± 1.3	59.2
Q392A	6.15 ± 0.21	24.5 ± 0.2	3.98	0.30 ± 0.02	111 ± 4	372
H103A	2.73 ± 0.11	15.6 ± 0.1	5.70	0.16 ± 0.01	34.5 ± 1.3	222

^a *Am*PDH produced in the natural source *Agaricus meleagris*; data from [2]

^b *Am*PDH recombinantly produced in *Pichia pastoris*; data from [3]

^c *Am*PDH with a His₆-tag, recombinantly produced in *Pichia pastoris*; this study

^d ND, not determined

Compared to *Am*PDH, the lowest increase (3.3-fold) in $K_{m,\text{Glc}}$ was observed for variant H103A, which lacks the covalent bond to FAD, and a reduction in $k_{\text{cat,Glc}}$ by a factor of only 2.7, resulting in the highest catalytic efficiency obtained for any variant ($k_{\text{cat}}/K_m = 5.7 \text{ mM}^{-1} \text{ s}^{-1}$). These results nicely fit to previous kinetic studies on *Am*PDH variant H103Y ($K_{m,\text{Glc}} = 3.85 \text{ mM}$; $k_{\text{cat,Glc}} = 27.4 \text{ s}^{-1}$) [18] as well as *Tm*P2O variant H167A ($K_{m,\text{Glc}} = 3.60 \text{ mM}$; $k_{\text{cat,Glc}} = 8.83 \text{ s}^{-1}$), which also lacks the covalent bond to the FAD [7]. Disrupting the covalent FAD linkage in H103A causes a decrease in flavin oxidative power (see section 5.3.7), as indicated by the decrease in the turnover number for GLC from 42.2 to 15.6 s^{-1} . Compared to other variants presented in the current study, His-103 is located on the *si*-side of the FAD and does not directly interact with the substrate. Therefore, mutations on this residue should not affect substrate binding to such an extent as for the other variants, reflected in the lowest $K_{m,\text{Glc}}$ value for H103A compared to other variants. Nevertheless, the 3.3-fold increase in $K_{m,\text{Glc}}$ for variant H103A indicates that covalent flavinylation is still important for catalysis and substrate binding, likely by positioning the flavin ring for optimal interaction. This was already observed for *Tm*P2O variant H167A (disrupted covalent bond to FAD), which showed a 5-fold increase in $K_{m,\text{Glc}}$ [7]. Therein, Kujawa and coworkers explained this observation with a 0.4–0.5 Å shift of the flavin dimethylbenzoid ring towards the substrate-binding site, which was observed in the crystal structures of wild-type *Tm*P2O but not of H167A. These observations are in line with the higher FAD mobility that was reported for *Am*PDH variant H103Y, lacking the covalent bond to FAD, by using electron paramagnetic resonance (EPR) spectroscopy [19]. This higher FAD mobility counteracts the fine-tuned positioning of the isoalloxazine ring with respect to the sugar substrate.

*Am*PDH variants Q392A and Y510A showed an 8-fold and 74-fold increase in $K_{m,\text{Glc}}$, respectively, confirming the importance of Gln-392 and Tyr-510 in substrate binding as proposed previously [15,17]. For Tyr-510, a more pronounced role was observed, which corroborates its proposed role to form the ‘floor’ below the sugar ring to ensure an optimal induced fit [17].

As mentioned earlier, variants V511F and V511W were generated based on an *in silico* study [15], which showed that the backbone oxygen atom of Val-511 is important for substrate binding. In the present study, we wanted to validate this hypothesis by introducing phenylalanine and tryptophan at position 511, both having bulky side chains that potentially disrupt the interaction between substrate and the backbone oxygen atom. For variant V511F, this nicely worked as indicated by a 130-fold increase in $K_{m,\text{Glc}}$ to 106 mM, which is in fact the highest K_m of all tested variants. A $k_{\text{cat,Glc}}$ of 19.9 s^{-1} (about 50% of wild-type *Am*PDH) indicates that once the substrate binds it is still rapidly oxidized. For variant V511W, a k_{cat} of 18.6 s^{-1} shows that GLC turnover is not dramatically decreased for this variant as well. Surprisingly, the $K_{m,\text{Glc}}$ for V511W increased only 4.5-fold, which is counterintuitive, as one would expect more interference in substrate binding by the bulkier tryptophan side chain, resulting in a higher $K_{m,\text{Glc}}$ for V511W compared to V511F. This behavior was rationalized by molecular dynamics (MD) simulations (**Figure 6**). The N-atom of the Trp-511 side chain forms two prominent H-bonds with the backbone O-atom of Asp-90 and Pro-92, cumulating to an average value of 0.75 observed H-bonds. This H-bonding causes the Trp-511

side chain to flip out of the active site. We think this is an excellent example how experiments and simulations can complement each other to provide a deeper understanding of the molecular basis of enzyme function.

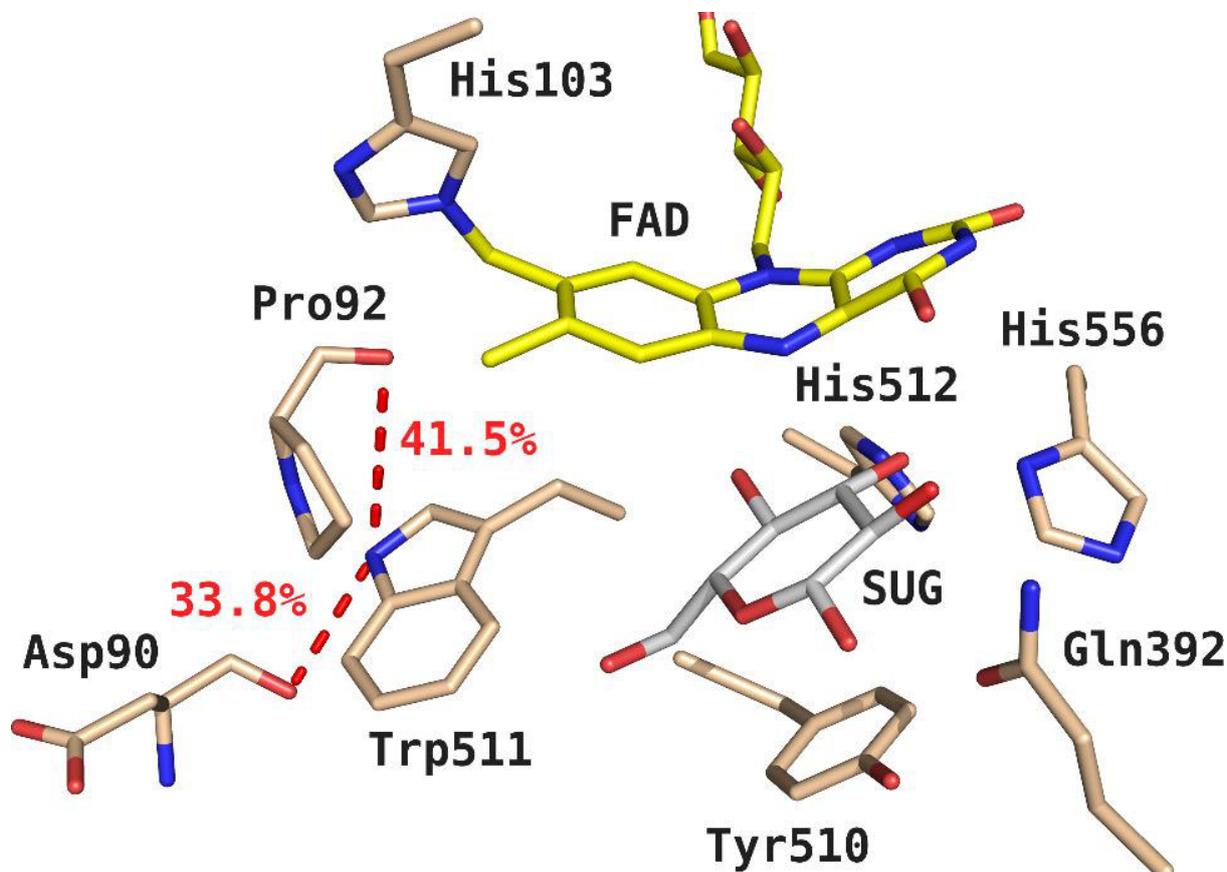


Figure 6: H-bonds formed between the side chain N-atom of Trp-511 and the backbone O-atom of Asp-90 and Pro-92 during the MD simulations of variant V511W. During the entire simulation, on average 0.75 H-bonds were observed between the mentioned atoms. Color code as in **Figure 1**.

To further support experimental results with MD simulations, relative binding free energies ($\Delta\Delta G_{\text{bind}}$) were estimated from experimental apparent K_m values ($\Delta\Delta G_{\text{bind}}^{\text{exp}}$) according to equation 1 and calculated from MD simulations ($\Delta\Delta G_{\text{bind}}^{\text{sim}}$) according to the linear interaction energy (LIE) method (equation 2). Because we wanted to investigate the effects on GLC binding, conversion, and product formation, only variants in which the mutated residues were expected to interact directly with GLC were subjected to MD simulations. As reported previously for *Am*PDH and other promiscuous enzymes [15,66], van der Waals interactions have been identified as the main driving force for substrate binding. These findings are corroborated in this study: the best match between $\Delta\Delta G_{\text{bind}}^{\text{exp}}$ and $\Delta\Delta G_{\text{bind}}^{\text{sim}}$ was obtained for the parameter set of $\alpha=0.503$ and $\beta=0$ [67] in equation 2, resulting in a root mean square error (RMSE) of only 3.9 kJ mol^{-1} , thereby neglecting the electrostatic β -term and using the van der Waals dependent α -term only.

Table 3: Relative binding free energies ($\Delta\Delta G_{bind}$) estimated from experimental apparent K_m values ($\Delta\Delta G_{bind}^{exp}$) according to equation 1, compared to $\Delta\Delta G_{bind}$ calculated from MD simulations ($\Delta\Delta G_{bind}^{sim}$) according to the linear interaction energy (LIE) method (equation 2). $\Delta\Delta G_{bind}^{sim}$ was calculated with $\alpha=0.503$ and $\beta=0$.

Variant	$\Delta\Delta G_{bind}$ [kJ/mol]	
	$\Delta\Delta G_{bind}^{exp}$	$\Delta\Delta G_{bind}^{sim}$
Q392A	5.1	5.6
Y510A	10.8	15.8
V511F	12.2	5.2
V511W	3.8	2.9
H512A	5.1	9.8
H556A	11.8	9.8
H556N	12.0	9.7

In **Table 3**, the difference between $\Delta\Delta G_{bind}^{exp}$ and $\Delta\Delta G_{bind}^{sim}$ is below the thermal noise $k_B T$ (2.5 kJ mol⁻¹) for Q392A, V511W, H556A, and H556N. For the remaining variants, the following differences in $\Delta\Delta G_{bind}^{exp}$ and $\Delta\Delta G_{bind}^{sim}$ were observed, which are slightly higher than the chemical accuracy [68,69] of 4 kJ mol⁻¹: 5 kJ mol⁻¹ (Y510A), 7 kJ mol⁻¹ (V511F), and 4.7 kJ mol⁻¹ (H512A). For Y510A, both $\Delta\Delta G_{bind}^{exp}$ and $\Delta\Delta G_{bind}^{sim}$ are very high, indicating a qualitative agreement. According to $\Delta\Delta G_{bind}^{exp}$ for V511F, the phenylalanine side chain in this variant seems to cause steric clashes that interfere with tight substrate binding, causing its faster dissociation. To simulate the unbinding, much longer simulation time scales would be necessary, which is beyond the scope of this study. For variant H512A, which lacks the catalytic His-512, the difference between $\Delta\Delta G_{bind}^{exp}$ and $\Delta\Delta G_{bind}^{sim}$ is possibly caused by a different reaction mechanism compared to *Am*PDH. The very low, but measurable residual activity of H512A could probably be derived from rarely occurring solvent-mediated proton abstraction of the GLC hydroxyl group attached to C2 or C3. This would allow for a hydride transfer to occur between the GLC H-atom attached to C2 or C3 and the flavin N5-atom. Therefore, we refrain from drawing any further conclusions based on the conducted MD simulations.

5.3.4 Mode of GLC oxidation: predictions from MD simulations verified by product analysis via GC-MS

We have measured the distances between (i) the GLC H-atom attached to C2 or C3 (HC2 and HC3) and the flavin N5-atom and (ii) the H-atom of the hydroxyl group of GLC attached to C2 or C3 (HO2 and HO3) and the His-512 N δ - or N ϵ -atom in the MD simulations. These distances were successfully utilized to reproduce and predict C2 and C3 oxidation of monosaccharides by *Am*PDH in recent studies [15,20], in which a 0.3-nm cutoff was used to discriminate between reactive (≤ 0.3 nm) and non-reactive distances (*i.e.* to enable proton abstraction and hydride

transfer). A qualitative agreement between the occurrences of reactive distances ≤ 0.3 nm and the experimentally determined $k_{\text{cat,Glc}}$ values is observed **Figure 7**.

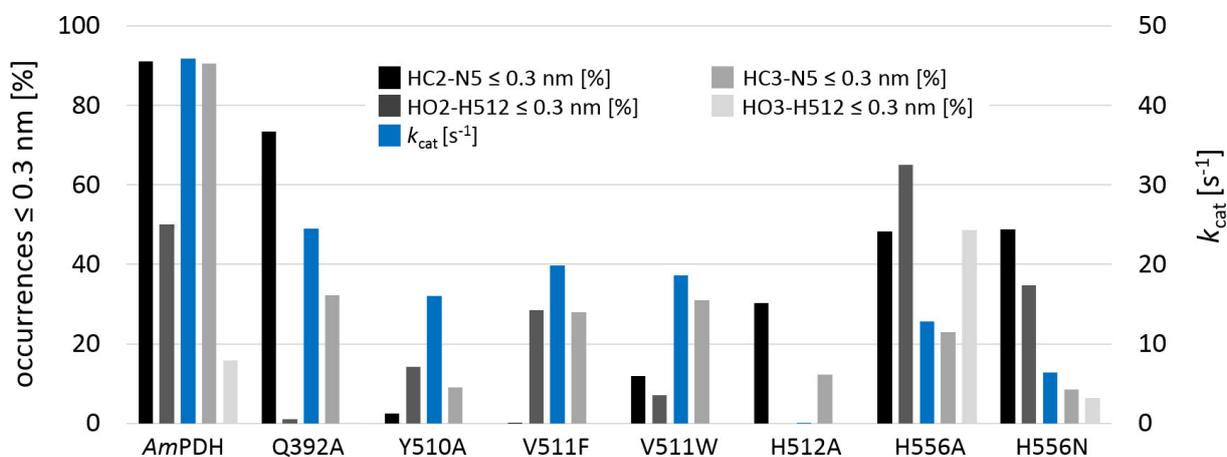


Figure 7: Distance distributions from MD simulations between (i) the D-glucose (GLC) H-atom attached to C2 or C3 (HC2 and HC3) and the flavin N5-atom and (ii) the GLC hydroxyl H-atom attached to C2 or C3 (HO2 and HO3) and the His-512 N_δ- or N_ε-atom, given in occurrences ≤ 0.3 nm [%]. The legend displays the color code for the corresponding distance. The distance distributions are compared to the experimentally derived k_{cat} values [s^{-1}] for GLC in blue bars (also see **Table 2**). Only variants directly interacting with GLC were considered for MD simulations, hence, H103A was not considered.

When multiplying the probability of occurrences of distances ≤ 0.3 nm for HC-N5 with those for HO-H512 (for C2 or C3, respectively), the product always yields a higher value for C2 compared to C3, indicating that C2 is more likely to be better positioned for the proton abstraction and hydride transfer than C3. This suggests that oxidation at C2 is preferred over C3. To verify this hypothesis, GC-MS measurements were conducted to quantify GLC and its (di-)keto-GLC reaction products. We used a novel approach by combining two-step derivatization (ethoxymation and trimethylsilylation), capillary gas chromatography, and electron ionization mass spectrometry (GC-EI-MS) to quantify GLC and its (di-)keto reaction products [70,71]. Since authentic standards for 3-keto-GLC and 2,3-diketo-GLC are not commercially available, capillary gas chromatography and high-resolution chemical ionization mass accuracy time of flight (GC-CI-QToFMS) was employed to identify them by their mass accuracy, isotopologue fraction pattern, and modified retention indices. For a detailed description of the performed GC-MS measurements, the reader is referred to materials and methods (section 5.3.10) as well as to **Figure S1A, S1B, and S1C** and **Table S2**. The final results obtained from GC-MS measurements are presented in **Table 4**, where the concentrations [mM] and ratios [%] of GLC and its (di-)keto reaction products are listed for AmPDH and its variants after 3, 8, and 120 min of reaction time. Only the reaction products of variant H512A could not be identified because of its extremely slow GLC turnover (see **Table 2**).

Table 4: GC-MS analysis of D-glucose (GLC) and its reaction products 2-keto-GLC, 3-keto-GLC, and 2,3-diketo-GLC. The reaction mixture contained 1 mL 10 mM potassium phosphate buffer pH 7.0, 1 U of the respective variant, 45 mM 1,4-benzoquinone as electron acceptor, and initially 16.7 mM GLC. The reaction was performed in a 1.5 mL vial at 30°C and 600 rpm in the dark. Samples (200 μ L) were retrieved after 3, 8, and 120 min reaction time, respectively. After deproteinizing through a 10-kDa ultrafiltration membrane, the filtrate was utilized for subsequent GC-MS analysis. Refer to the materials and methods, section 5.3.10 for a detailed description of the sample preparation procedure for the GC-MS measurements. For GLC and 2-keto-GLC, an external calibration was possible with commercially available standards, allowing for a quantitative determination of their concentrations (in mM) with a detection limit of 0.2 μ M. Because no standards for 3-keto-GLC and 2,3-diketo-GLC were commercially available, only semiquantitative results (relative concentrations in mM) could be calculated, based on the peak-area-ratios of the respective analyte to 2-keto-GLC at a given mass to charge ratio. To obtain the 'Ratio', the percentage of each compound with respect to the sum of (relative) concentrations was calculated for each variant.

Variant	Reaction time [min]	Concentration [mM]		Relative concentration [mM]		Ratio [%]			
		GLC	2-keto-GLC	3-keto-GLC	2,3-diketo-GLC	GLC	2-keto-GLC	3-keto-GLC	2,3-diketo-GLC
AmPDH	3	2.689	1.720	0.000	0.061	60.2	38.5	0.0	1.4
	8	2.432	1.686	0.001	0.273	55.4	38.4	0.0	6.2
	120	0.427	1.331	0.026	1.271	14.0	43.6	0.9	41.6
H103A	3	3.417	0.597	0.000	0.100	83.1	14.5	0.0	2.4
	8	2.906	1.652	0.004	0.342	59.3	33.7	0.1	7.0
	120	1.624	1.491	0.024	1.108	38.2	35.1	0.6	26.1
Q392A	3	3.072	1.503	0.002	0.186	64.5	31.6	0.0	3.9
	8	2.626	1.870	0.002	0.724	50.3	35.8	0.0	13.9
	120	1.041	1.485	0.032	1.540	25.4	36.2	0.8	37.6
Y510A	3	3.353	0.126	0.000	0.002	96.3	3.6	0.0	0.1
	8	3.401	0.386	0.000	0.034	89.0	10.1	0.0	0.9
	120	2.732	1.590	0.010	0.401	57.7	33.6	0.2	8.5
V511F	3	3.617	0.311	0.001	0.005	91.9	7.9	0.0	0.1
	8	3.604	1.090	0.001	0.030	76.3	23.1	0.0	0.6
	120	2.420	2.464	0.025	0.387	45.7	46.5	0.5	7.3
V511W	3	3.335	1.701	0.001	0.058	65.5	33.4	0.0	1.1
	8	3.037	2.458	0.010	0.331	52.0	42.1	0.2	5.7
	120	0.275	1.835	0.150	1.564	7.2	48.0	3.9	40.9
H556A	3	3.855	0.293	0.001	0.018	92.5	7.0	0.0	0.4
	8	3.464	0.811	0.000	0.075	79.6	18.6	0.0	1.7
	120	3.538	2.804	0.001	0.466	52.0	41.2	0.0	6.8
H556N	3	3.922	0.090	0.000	0.003	97.7	2.2	0.0	0.1
	8	3.656	0.187	0.000	0.012	94.8	4.9	0.0	0.3
	120	3.447	1.169	0.001	0.110	72.9	24.7	0.0	2.3

The results in **Table 4** support the hypothesis derived from MD simulations that C2 oxidation is indeed preferred over C3 oxidation, since only very little or no 3-keto-GLC was detected but significant amounts of 2-keto-GLC as well as 2,3-diketo-GLC for all tested variants. This indicates that 2-oxidation is significantly faster than 3-oxidation, both for GLC as well as for a mono-oxidized intermediate. These results nicely agree with previous studies employing fluorinated GLC derivatives, where 3-fluoro-3-deoxy-GLC (3FG) was rapidly oxidized, whereas 2-fluoro-2-deoxy-GLC (2FG) only very slowly [17]. Based on these results, we propose here a scheme for the two-step oxidation of GLC by *Am*PDH that is also valid for all tested variants (**Figure 8**). This scheme is very similar to the one reported earlier by Volc and coworkers for PDH from *Agaricus bisporus* (*Ab*PDH) [72]. The main difference between *Ab*PDH and *Am*PDH is that *Am*PDH almost exclusively dioxidizes GLC *via* 2-keto-GLC, while *Ab*PDH performs this reaction *via* both 2-keto-GLC and 3-keto-GLC, with the latter being favored to some extent, to yield the final product 2,3-diketo-GLC.

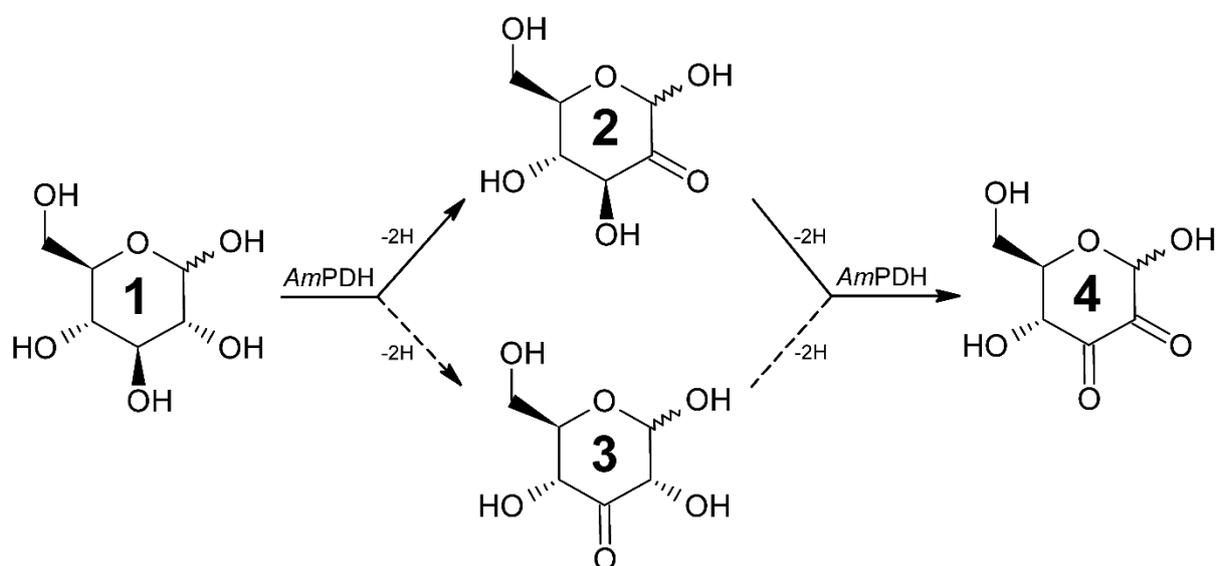


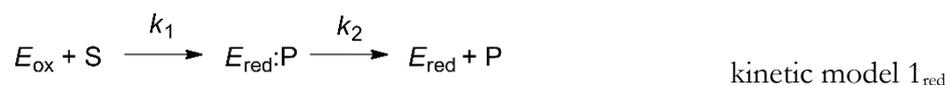
Figure 8: Proposed reaction scheme for the two-step oxidation of D-glucose (GLC) by *Am*PDH and the investigated variants according to GC-MS data (**Table 4**) and molecular dynamics simulations (**Figure 7**). Dioxidation of GLC (compound **1**) to 2,3-diketo-GLC (compound **4**) proceeds predominantly via 2-keto-GLC (compound **2**), which is slowly oxidized at C3 to form the final product 2,3-diketo-GLC (route indicated by solid arrows). Only a very small portion of GLC is initially oxidized to 3-keto-GLC (compound **3**), which is rapidly converted to 2,3-diketo-GLC (route indicated by dashed arrows). This reaction scheme is very similar to the one proposed for *Agaricus bisporus* PDH (*Ab*PDH) by Volc and coworkers [72]. However, 2-keto-GLC and 3-keto-GLC formation were reported to be equally favored by *Ab*PDH, which are finally both converted to 2,3-diketo-GLC.

In a previous *in silico* study, mutations at Gln-392, Tyr-510, and Val-511 were suggested to affect the preferred GLC oxidation site, as these amino acids were involved in prominent H-bonds with the substrate bound in poses relevant for either C2 or C3 oxidation [15]. Interestingly, neither explicit simulations of the mutants, nor the GC-MS experiments indicated an altered preference

for C2 or C3 oxidation compared to *Am*PDH. Consequently, substrate binding in an orientation specific for C2 or C3 oxidation cannot be attributed to one specific active site amino acid in *Am*PDH only, but an interplay between various residues seems to be critical for this.

5.3.5 Preference of *Am*PDH for C2 or C3 oxidized substrates

The reductive half reaction ($\text{FAD} \rightarrow \text{FADH}_2$) of *Am*PDH with substrates that are specifically oxidized at either C2 or C3 was investigated to test the preferred mode of sugar oxidation of *Am*PDH. As reported previously, *Am*PDH is reduced extremely fast by GLC at 30°C, with an apparent bimolecular rate constant ($k_{\text{app},30^\circ\text{C}}$) of $1.1 \pm 0.03 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ [19]. Consequently, all stopped-flow experiments in the present study were conducted at 4°C to slow down the reaction and capture as much kinetic information as possible. All obtained kinetic traces are plotted in **Figure S2** and were simulated with parameters listed in **Table S3** according to kinetic model 1_{red}:



Refer to materials and methods (section 5.2.11) for more details on the simulations of kinetic traces. According to previously published data [1,14], D-galactose and L-arabinose are exclusively oxidized at C2, whereas methyl- α -D-glucopyranoside is only oxidized at C3, with no double oxidation occurring. *Am*PDH oxidizes GLC at both C2 and C3, but C3 oxidation – also of mono-oxidized 2-keto-GLC – is very slow and takes minutes to hours (section 5.3.4, **Table 4**, and **Figure 8**). Therefore, GLC can be treated as C2 oxidized substrate in the time-course of stopped-flow experiments that spans just a few seconds. The reactions of *Am*PDH for all investigated electron donors (sugars) were monitored at 463 nm and were biphasic and their time traces were fitted by a double-exponential function to obtain the k_{obs} value for the corresponding sugar concentration. The first phase was the FAD reduction with a large absorbance decrease at 463 nm and was used for calculating $k_{\text{app},4^\circ\text{C}}$ (**Table 5**). The second phase showed a small decrease in absorbance at the same wavelength and was independent of the GLC concentration (data not shown). The y-intercepts between 7.6-15.2 s⁻¹ indicate a reversible step, which cannot be significant, though: on the one hand, the magnitude of the y-intercept is very small compared with the scale of the y-axis for all tested sugars (**Figure S2**). Moreover, the simulated traces according to kinetic model 1_{red} nicely agree with experimental traces (**Figure S2**), which indicates the validity of the utilized model, which does not consider a reversible step. For all sugars, the obtained $k_{\text{app},4^\circ\text{C}}$ were very similar and in a range of 4.4 to $6.8 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$. The pentose L-arabinose had the highest $k_{\text{app},4^\circ\text{C}}$, which can possibly be attributed to its smaller size compared to the other tested sugars (hexoses), which could ultimately facilitate its diffusion to and from the active site. Based on the obtained $k_{\text{app},4^\circ\text{C}}$ values we conclude that *Am*PDH equally favors C2 and C3 oxidized substrates.

Table 5: Stopped-flow spectroscopy to monitor the bimolecular rate constants for the reductive half-reaction of *AmPDH* with D-glucose (GLC), D-galactose (GAL), L-arabinose (ARA), and methyl- α -D-glucopyranoside (MGP) as well as the oxidative half-reaction with 1,4-benzoquinone (BQ). For variant H556A, only GLC and BQ were used. Experiments were performed in 50 mM potassium phosphate buffer pH 7.0 and 4°C. Stopped-flow traces were simulated with the program Berkeley Madonna (version 8.3.14) according to kinetic model 1_{red} or kinetic model 1_{ox} (refer to materials and methods for further information). Employed enzyme and substrate concentrations, monitored wavelengths, as well as experimental and simulated traces can be found in **Figures S2–S4**. More details on GC-MS measurements to determine the GLC oxidation sites for H556A can be found in **Table 4**.

Variant	Substrate	Oxidation site	Reference for oxidation site	Apparent bimolecular rate constant k_{app} [M ⁻¹ s ⁻¹]	
				Experiment (y-intercept)	Simulation
<i>AmPDH</i>	GLC	C2/3 ^a	[10,14]	$5.4 \pm 0.11 \times 10^4$ (7.61)	6.9×10^4
	GAL	C2	[14]	$4.4 \pm 0.15 \times 10^4$ (9.11)	5.2×10^4
	ARA			$6.8 \pm 0.10 \times 10^4$ (9.06)	7.8×10^4
	MGP	C3	[11]	$4.8 \pm 0.05 \times 10^4$ (15.2)	6.5×10^4
	BQ	-	-	$6.6 \pm 0.17 \times 10^3$ (0.92)	8.0×10^3
H556A	GLC	C2/3 ^a	this work	$7.3 \pm 0.24 \times 10^1$ (0.06)	8.0×10^1
	BQ	-	-	$4.1 \pm 0.02 \times 10^4$ (-0.46)	3.6×10^4

^a Predominantly oxidized at C2 and only very slowly at C3 as reported in section 5.3.4 and **Table 4**

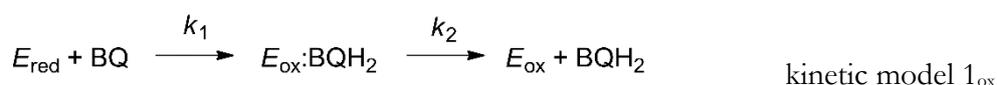
5.3.6 Effects of the His → Ala replacement at position 556 on both half reactions

Although His-556 does not act as catalytic base, we found that it is important for sugar substrate binding (see section 5.3.3). To identify the rate-limiting step in variant H556A and compare the results to recombinant wild-type *AmPDH*, the reductive- (FAD → FADH₂) and oxidative (FADH₂ → FAD) half-reactions of these enzymes were investigated by stopped-flow spectroscopy. The reductive half-reactions were performed with D-glucose (GLC) as electron donor and the oxidative half-reactions with the artificial electron acceptor 1,4-benzoquinone (BQ). In a previous study, the ferrocenium ion (Fc⁺) was used as electron acceptor, which has an absorbance at 330–350 nm and 625 nm [19]. Therefore, Fc⁺ interferes to a greater extent with the flavin absorbance than BQ, which has absorption maxima at 244 nm, and the corresponding hydroquinone at 222 nm and 296 nm, respectively [73]. The stopped-flow experiments were essentially performed as described in section 5.3.5 and the kinetic were recorded at 465 nm for H556A and at 463 nm for *AmPDH*, respectively. For both, GLC and BQ, the reactions were biphasic and their time traces were fitted by a double-exponential function to obtain the $k_{obs,4^\circ C}$ value.

For details on the reductive half-reaction of *AmPDH* with GLC, refer to section 5.3.5, **Figure S2**, and **Table 5**, for which a $k_{app,4^\circ C}$ of $5.4 \pm 0.11 \times 10^4$ M⁻¹ s⁻¹ was obtained, corresponding to a 2-fold decrease compared to measurements at 30°C [19]. The reaction rate approximately doubles when

the temperature is increased by 10°C, however, this strongly depends on the activation energy of the reaction [74,75]. Therefore, the agreement with the previously published $k_{app,30^\circ C}$ of *Am*PDH with GLC is good.

For the oxidative half-reaction of *Am*PDH with BQ, the first phase of the fit corresponds to FAD oxidation with a large absorbance increase at 463 nm (**Figure S3**), which was used to calculate $k_{app,4^\circ C}$ (**Table 5**). The second phase showed a significant absorbance decrease from the intermediate species $E_{ox}:BQH_2$ to E_{ox} and was independent of the BQ concentration (data not shown). For the first phase, the y-intercept is very small (0.92 s^{-1}), indicating a negligible reversible step. Simulations according to kinetic model 1_{ox}



with parameter sets listed in **Table S3** corroborate this observation, as the simulated traces nicely agree with experimental traces (**Figure S3**) and the model does not consider a reversible step as well. The second phase of simulated and experimental traces does not overlap as nicely as the first phase, which can be attributed to the complicated quinone oxidation process or to polymerizations in side reactions. However, the imperfect fit does not affect the first phase, which was used to calculate k_{obs} for *Am*PDH oxidation, and therefore, the utilized kinetic model is valid.

For the reductive half-reaction of H556A with GLC, the second phase of the fit corresponds to FAD reduction with a large absorbance decrease at 465 nm (**Figure S4**), which was used to calculate $k_{app,4^\circ C}$ (**Table 5**). The first phase only showed a small absorbance decrease and was independent of the GLC concentration (data not shown). For the second phase, the y-intercept is close to zero (0.06 sec^{-1}), indicating a negligible reversible step. Simulations according to kinetic model 1_{red} (see section 5.3.5 for more details on the model) with parameter sets listed in **Table S3** corroborate this observation, as the simulated traces nicely agree with experimental traces (**Figure S4**) and the model does not consider a reversible step. The $k_{app,4^\circ C}$ for the reductive half-reaction of variant H556A with GLC dramatically decreased 720-fold when compared to *Am*PDH (**Figure S4** and **Table 5**), indicating that the two-electron transfer to FAD is significantly hampered by this amino acid exchange.

In the oxidative half reaction of H556A with BQ, the first phase of the fit corresponds to FAD oxidation with a large absorbance increase at 465 nm (**Figure S4**), which was used to calculate $k_{app,4^\circ C}$ (**Table 5**). The second phase only showed a small absorbance increase and was independent of the BQ concentration (data not shown). For the first phase, the y-intercept is again close to zero (-0.46 sec^{-1}), which indicates no reversible step. Simulated traces according to kinetic model 1_{ox} with parameter sets listed in **Table S3** agree well with experimental traces (**Figure S4**), which corroborates the lack of a reversible step. Similar as for *Am*PDH, the second phase of simulated and experimental traces for the oxidative half-reaction in H556A do not perfectly overlap, which can be attributed to the complicated quinone oxidation process or to polymerizations in side

reactions. However, the imperfect fit does not affect the first phase, which was used to calculate k_{obs} for H556A oxidation, and therefore, the utilized kinetic model is valid. For variant H556A, the $k_{\text{app},4^{\circ}\text{C}}$ of the biphasic oxidative half-reaction with BQ was about 6.2-fold higher compared to *Am*PDH (**Table 5**), which can possibly be attributed to the lower redox potential of H556A (see section 5.3.7) or to a more accessible active-site in the variant. However, since the binding site of BQ in *Am*PDH and the electron transfer path from *Am*PDH to BQ are still elusive, further studies will be necessary to draw final conclusions.

The presented pre-steady-state experiments clearly demonstrate that the drastically impaired reductive half-reaction of variant H556A is the main reason for its much lower catalytic efficiency (**Table 2**) compared to *Am*PDH. The oxidative half-reaction of variant H556A is significantly faster compared to *Am*PDH.

5.3.7 Redox potential and oxygen reactivity

Covalent FAD attachment in flavoproteins has been extensively studied and is known to increase the redox properties of the enzyme [26,76], leaving only high-potential electron acceptors such as molecular oxygen as redox partners (see section 5.3.2). Consequently, a covalently attached FAD was mainly found in oxidases. Furthermore, electron-withdrawing substituents at position 8 substantially increase the flavin redox potential [77]. This makes *Am*PDH a rather unusual example for a dehydrogenase in carrying a monocovalently bound FAD *via* position 8. Therefore, we probed the effect of the missing covalent FAD linkage in H103A on its redox potential. A drastic decrease in redox potential for variant H556A could potentially rationalize the 720-fold decrease of its reductive half-reaction with D-glucose (see section 5.3.6), which is why we determined its redox potential as well. In this study, Massey's method [27] employing the xanthine and xanthine oxidase reduction system was used for the first time to determine the redox potentials of *Am*PDH, H103A, and H556A, which were subsequently to each other (**Table 6** and **Figure 9**).

Table 6: Redox potential values (E_m^0). Measurements and calculations were performed as described in the materials and methods (section 5.2.12). Reduction-spectra are plotted in **Figure 9**. A slope close to 1.0 indicates a two-electron reduction step, which is why no one-electron reduction steps ($E_1^{0'}$ and $E_2^{0'}$) could be determined for *Am*PDH. A slope close to 0.5 indicates a one-electron reduction step, which was observed for variant H103A and H556A. Therefore, both one-electron reduction steps could be determined for these variants. *M* indicates the maximum fraction of thermodynamically stable semiquinone formed, which was necessary to calculate E_m^0 for both variants.

Variant	E_m^0 [mV]	$E_1^{0'}$ [mV]	$E_2^{0'}$ [mV]	Slope	<i>M</i>
<i>Am</i> PDH	+92	-	-	0.97	-
H103A	+63	+113	+13	0.41	0.78
H556A	+84	+129	+39	0.50	0.74

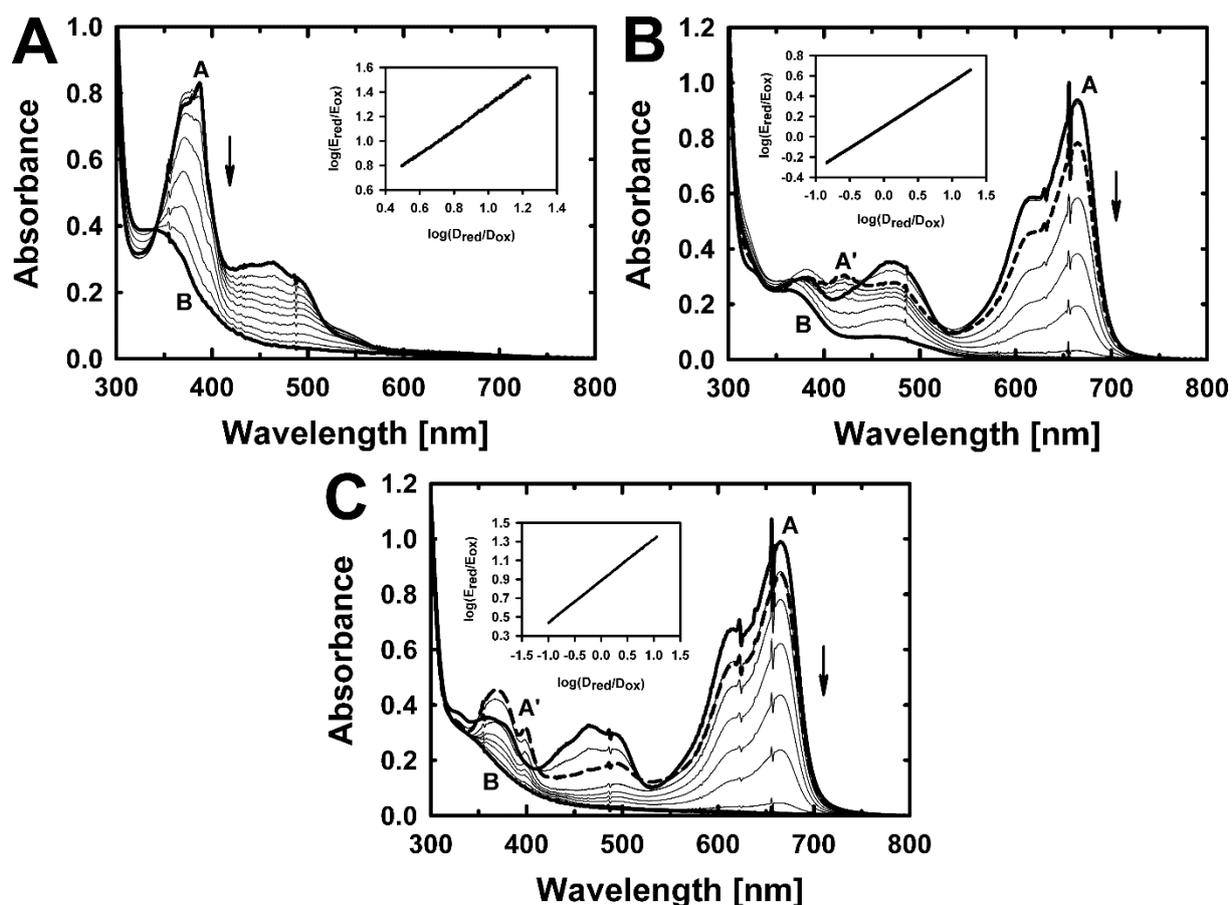


Figure 9: Determination of reduction potentials according to Massey's method [27] for AmPDH (A), variant H103A (B), and variant H556A (C). In each panel, spectral line A (uppermost solid line) represents the fully oxidized enzyme and dye, whereas spectral line B (lowest solid line) represents fully reduced enzyme and dye. In panel B and C, dashed line A' represents the maximum of an occurring semiquinone intermediate with a maximum absorbance at 422 nm (B) or 368 nm (C). The inset in panel A shows the reduction of phenazine methosulfate (dye, 'D') monitored at 387 nm and the reduction of AmPDH (enzyme, 'E') monitored at 463 nm, allowing the calculation of the ratios between reduced and oxidized species (E_{red}/E_{ox} and D_{red}/D_{ox}). The reduction potential of AmPDH was calculated to be $+92 \pm 3$ mV. The inset in panel B shows the reduction of methylene blue ('D') monitored at 664 nm and the reduction of variant H103A ('E') monitored at 460 nm. The reduction potential of variant H103A was calculated to be $+63 \pm 3$ mV. The inset in panel C shows the reduction of methylene blue ('D') monitored at 664 nm and the reduction of variant H556A ('E') monitored at 465 nm. The reduction potential of variant H556A was calculated to be $+84 \pm 4$ mV. An overview of the obtained data is given in **Table 6**.

The reduction of AmPDH (**Table 6** and **Figure 9A**) followed a two-electron reduction process as judged by a slope of 1.0 when plotting $\log(E_{red}/E_{ox})$ versus $\log(D_{red}/D_{ox})$. The midpoint potential (E_m^0) for AmPDH was $+92 \pm 3$ mV at pH 7.0. According to the equation $\Delta E_m/\Delta pH = -0.06$ V [52,53], the E_m^0 value at pH 7.4 was calculated to be $+116 \pm 3$ mV. Considering the different methodological approaches and the measurement uncertainty, this agrees well with previously

published E_m^o values for *Am*PDH, which were $+135 \pm 10$ mV [19] or about $+150$ mV [78], recorded spectroelectrochemically at pH 7.4 and 25°C.

For H103A (**Table 6** and **Figure 9B**), the fully oxidized enzyme was reduced by two one-electron reduction steps, as indicated by a slope of 0.41 when plotting $\log(E_{\text{red}}/E_{\text{ox}})$ versus $\log(D_{\text{red}}/D_{\text{ox}})$. Thereby, the reduction of H103A proceeds *via* a putative semiquinone intermediate to the fully reduced enzyme and the E_m^o of the overall reduction was calculated according to equation 5 and equation 6. The observed semiquinone intermediate for variant H103A had an absorption maximum at 422 nm (dashed line in **Figure 9B**). This intermediate was observed previously for variant H103Y (absorption maximum at 430 nm) and was found to have no physiological relevance [19]. This wavelength does not perfectly fit to the characteristic absorbance of the anionic semiquinone radical at around 380 and 400 nm [25]. However, it is known that the microenvironment around the flavin (*e.g.* the protein moiety, pH, ligands) can significantly affect the occurrence of semiquinone radicals [25] and can cause bathochromic shifts of *e.g.* the 450 nm flavin absorption band [64]. Consequently, the red-shifted semiquinone absorbance in variant H103A can possibly attributed to similar factors. For variant H103A, E_m^o was $+63 \pm 3$ mV at pH 7.0 and 25°C (**Table 6**), corresponding to $+87 \pm 3$ mV at pH 7.4. Therefore, disruption of the covalent FAD linkage decreases the redox potential by 29 mV. For *Tm*P2O, a very similar observation was made for variant H167A, which also lacks the covalent bond to FAD. The decrease in E_m^o was 32 mV for the His₆-tagged enzymes and 45 mV for the non-tagged enzymes when compared to the wild-type [7]. Disruption of the covalent FAD linkage in *Am*PDH variant H103Y caused a much more pronounced decrease in E_m^o of about 100 mV [19] and a 5.7°C lower thermal stability (section 5.3.2). One might speculate that the bulky tyrosine side chain destabilizes the active site sterically, which causes a significant decrease in E_m^o and T_m , whereas for H103A these effects are much less pronounced.

For H556A (**Table 6** and **Figure 9C**), similar observations as for H103A were made: the fully oxidized enzyme was reduced by two one-electron reduction steps, as indicated by a slope of 0.50 when plotting $\log(E_{\text{red}}/E_{\text{ox}})$ versus $\log(D_{\text{red}}/D_{\text{ox}})$. This again points towards a putative semiquinone intermediate on the way to the fully reduced enzyme and the E_m^o of the overall reduction was calculated according to equation 5 and equation 6. The observed intermediate for variant H556A had an absorption maximum at 368 nm (dashed line in **Figure 9C**) and was detected previously in the resting state of *Am*PDH, which could be attributed to an equilibrium of the oxidized form of the enzyme with its reduced forms [19]. For variant H556A, E_m^o was $+84 \pm 4$ mV at pH 7.0 and 25°C (**Table 6**), corresponding to $+108 \pm 4$ mV at pH 7.4. Therefore, the redox potential of variant H556A is only decreased by 8 mV compared to *Am*PDH, indicating that the difference in redox potential between H556A and *Am*PDH is not responsible for the 720-fold decrease of the reductive half-reaction with D-glucose (see section 5.3.6).

As mentioned earlier, it is not exactly known to date why flavoenzymes do or do not react with oxygen [18,28]. Factors such as a positive charge in the active site [79], its accessibility *via* channels

and tunnels [80], and the microenvironment around the flavin N5-C(4a) [29,30] have been identified as important features for oxygen reactivity. In a recent study with *Am*PDH variant H103Y, it was concluded that further studies on the oxygen reactivity of *Am*PDH are required to understand this feature in greater detail. Therefore, all generated variants were routinely probed for their steady-state oxygen reactivity, with variant H512A being the only exception, because it reacted too slowly with the electron donor GLC (see section 5.3.3). Results for relative activities are given in **Figure 10**. *Am*PDH had a very low oxygen reactivity of $0.104 \mu\text{M min}^{-1} \text{mg}^{-1}$ in absolute values. This is also in line with the slow reoxidation of reduced *Am*PDH by oxygen at air saturation in the course of many hours ($\sim 40\%$ reoxidation after 15 h) [17].

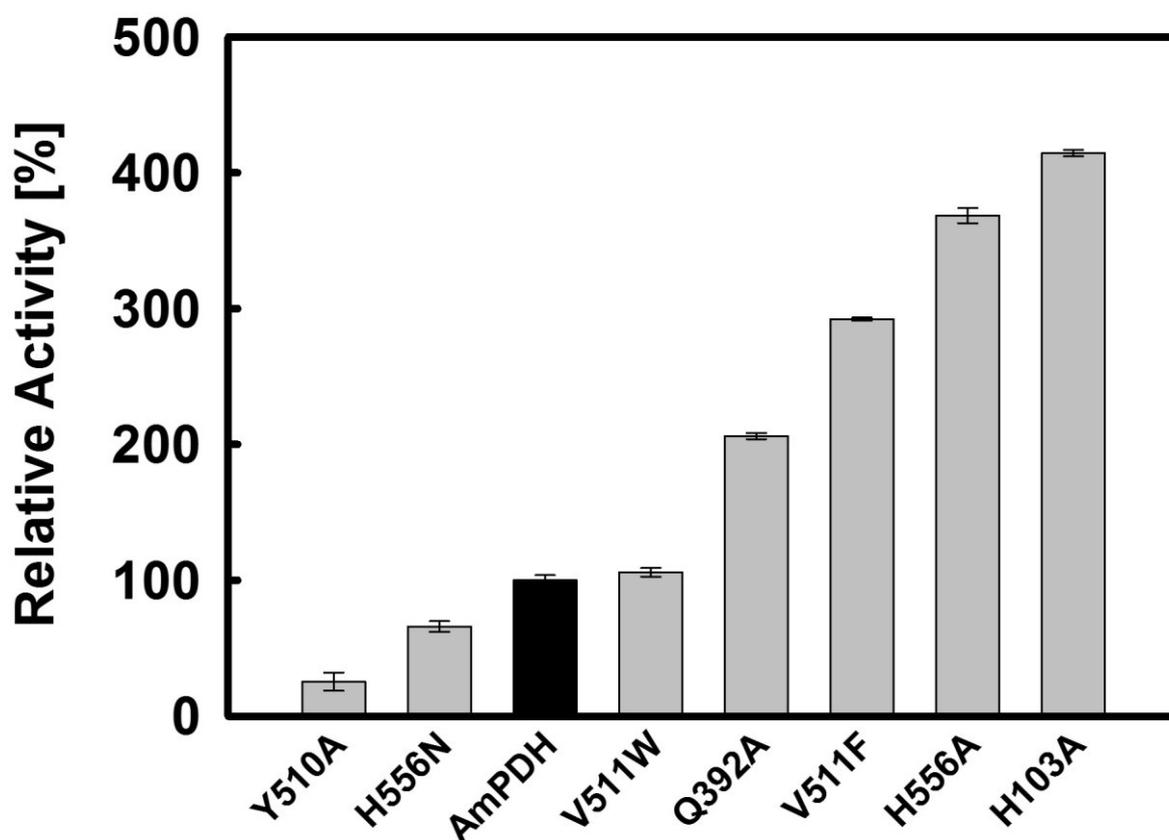


Figure 10: Relative oxygen reactivities. Steady-state oxygen reactivity was determined using an Amplex Red/horseradish peroxidase based assay as reported previously [18]. The assay contained 0.5 mg mL^{-1} of the purified enzyme, $50 \mu\text{M}$ Amplex Red, 25 mM D-glucose, 0.1 U mL^{-1} horseradish peroxidase and 50 mM sodium phosphate buffer pH 7.4. Error bars represent the standard deviation of four repeats.

Interestingly, the oxygen reactivity of only two out of seven tested variants was lower than for wild-type *Am*PDH: Y510A ($0.026 \mu\text{M min}^{-1} \text{mg}^{-1}$; 25% of *Am*PDH) and H556N ($0.068 \mu\text{M min}^{-1} \text{mg}^{-1}$; 65% of *Am*PDH). For Y510A, a possible reason could be the removal of the hydrophobic floor of the active site, which might make positioning of oxygen at a correct distance from the flavin

C(4a) problematic. This effect has previously been observed for aryl-alcohol oxidase variant F501W, that showed ~2-fold increase in oxygen reactivity compared to the wild-type [81].

Variant H556N was initially generated to emulate the catalytic pair of the oxidase *TmP2O*, consisting of His-548 (corresponding to His-512 in *AmPDH*) and Asn-593 (corresponding to His-556 in *AmPDH*), to eventually increase oxygen reactivity in *AmPDH*. Variant H556A showed the second highest increase in oxygen reactivity of all studied variants ($0.382 \mu\text{M min}^{-1} \text{mg}^{-1}$; 369% of *AmPDH*), however, this level is still far from a true oxidase. The results for the two His-556 variants supports the hypothesis that the shorter distance (3.5 Å) between the catalytic pair His-512 (Nε2) and His-556 (Nδ1) in *AmPDH* compared to 4.2 Å for His-558 (Nε2) and Asn-593 (Nδ2) in *TmP2O* could abolish C(4a) adduct formation or provide too strong stabilization thereof, which ultimately prevents oxygen reactivity [17].

The tryptophan side chain in variant V511W has essentially no effect on oxygen reactivity ($0.110 \mu\text{M min}^{-1} \text{mg}^{-1}$; 106% of *AmPDH*). Interestingly, the oxygen reactivity for variant V511F almost triples compared to *AmPDH* ($0.303 \mu\text{M min}^{-1} \text{mg}^{-1}$; 292% of *AmPDH*). For both, V511F and V511W, bulky, unpolar, and uncharged residues were introduced. This observation is in line with a previous study for choline oxidase (CHO), where a nonpolar side chain was found to be essential for guiding oxygen to the proximity of the flavin C(4a) [82].

The oxygen reactivity of H103A, which lacks the covalent bond to FAD, was $0.430 \mu\text{M min}^{-1} \text{mg}^{-1}$, corresponding to 414% of *AmPDH*. This is the highest oxygen reactivity amongst all tested variants in this study and in line with previously published work on variant H103Y, which was found to have a 5.3-fold higher oxygen reactivity compared to *AmPDH* [18]. Both variants (H103A and H103Y) lack the covalent bond to FAD and showed a significantly decrease in redox potential, which was more pronounced for H103Y [19] and both variants showed the highest oxygen reactivities of all tested *AmPDH* variants to date, which was again higher for H103Y [18]. The oxygen reactivity for variant H556A was increased to a similar level as for H103A, however, its redox potential was only slightly decreased compared to *AmPDH*. Consequently, a direct link between lower redox potential and higher oxygen reactivity cannot be drawn in *AmPDH*. Moreover, one would expect rather the opposite correlation – a high redox potential that is linked to increased oxygen reactivity, because a higher redox potential is usually associated with electron accepting redox-partners that have high redox potentials themselves, such as oxygen. It is therefore surprising that for *AmPDH* the opposite behavior is observed. On the other hand, this gives strong evidence that other factors play an important role as well, such as the microenvironment of the active site, or tunnels and channels guiding oxygen diffusion, and that oxygen reactivity has to be seen in the whole context of the respective flavoenzyme. Pre-steady-state measurements on the reoxidation rate of fully reduced enzymes with 0.13 mM O₂ gave the following k_{obs} values (graphs not shown): $2.6 \times 10^{-5} \text{ s}^{-1}$ for *AmPDH*, $1.2 \times 10^{-4} \text{ s}^{-1}$ for H103A, and $7.4 \times 10^{-4} \text{ s}^{-1}$ for H556A. These results strongly corroborate the hypothesis that in *AmPDH*, His-556 closely interacts with

(i) His-512 and (ii) with a presumable C(4a) adduct, which would ultimately interfere with oxygen reactivity [17].

The results on the redox potential and oxygen reactivity presented in this section provide valuable insights concerning the covalent attachment of the FAD and the amino acids in its vicinity. Based on these findings, we propose further studies to elucidate the mechanism of oxygen reactivity in *Am*PDH. i) We encourage creating multiple-variants, with combinations of the four mutations yielding the highest increase in oxygen reactivity (H103A, Q392A, Y510A, and V511F). Although H103Y has slightly higher oxygen activity compared to H103A [18], the thermostability of H103Y [19] is 5.7°C lower compared to H103A and lower than for all tested variants in the present study. Therefore, multiple-variants containing H103Y might rather be unstable. ii) Crystal structures of variants H103A, H556A, V511F, and Q392A have the potential to provide a detailed understanding about structural features governing oxygen reactivity. iii) Molecular dynamics (MD) simulations with structures of *Am*PDH and the variants suggested in point ii) could prove to be a powerful tool to elucidate oxygen diffusion pathways and reactivity. iv) Detection of the C(4a) adduct under physiological reaction conditions would prove that it is not only an X-ray radiation artifact.

5.4 Conclusion

In the current study, *Am*PDH and eight variants thereof were investigated by biochemical, biophysical and computational means. The secondary structures as well as the thermostability were not affected by the newly introduced amino acids as judged by ECD spectroscopy and *Thermo*FAD experiments, respectively. All enzymes were monocovalently flavinylated, only variant H103A had a non-covalently but tightly bound FAD cofactor. The redox potential for this variant was decreased by 29 mV, which is in line with other flavin-dependent oxidoreductases [26]. Variant H556A showed only a small decrease in redox potential of 8 mV. Apparent steady-state kinetics unequivocally demonstrated that His-512 is the sole catalytic base in *Am*PDH, and that His-556 is mainly relevant for GLC binding. Gln-392, Tyr-510, and Val-511 were identified to have important roles in substrate binding as proposed previously [15]. MD simulations and free energy calculations reproduced the experimental $K_{m, \text{Glc}}$ values well and helped to explain the counterintuitive observation that V511F has a much higher $K_{m, \text{Glc}}$ value compared to V511W. Predictions from MD simulations, which were verified by GC-MS measurements, unequivocally demonstrated that *Am*PDH and its variants preferentially oxidize GLC *via* 2-keto-GLC to the final product 2,3-diketo-GLC. Pre-steady-state kinetics demonstrated that *Am*PDH equally favors C2 or C3 oxidized substrates and that the reductive-half reaction in variant H556A is decreased 720-fold compared to *Am*PDH, which is the reason for its lower catalytic efficiency.

Compared to *Am*PDH, a decreased oxygen reactivity was only observed for variants Y510A and H556N, whereas all other variants had the same or significantly higher oxygen reactivity. We could find evidence for the hypothesis [17] that the short distance between the active site histidines His-512 and His-556 in *Am*PDH could abolish C(4a) adduct formation or stabilize it too much, thereby preventing oxygen reactivity. Based on our results, we propose new experiments to furthermore elucidate the oxygen reactivity in *Am*PDH.

5.5 References

- 1 Peterbauer CK & Volc J (2010) Pyranose dehydrogenases: Biochemical features and perspectives of technological applications. *Appl. Microbiol. Biotechnol.* **85**, 837–848.
- 2 Sygmund C, Kittl R, Volc J, Halada P, Kubátová E, Haltrich D & Peterbauer CK (2008) Characterization of pyranose dehydrogenase from *Agaricus meleagris* and its application in the C-2 specific conversion of D-galactose. *J. Biotechnol.* **133**, 334–342.
- 3 Sygmund C, Gutmann A, Krondorfer I, Kujawa M, Glieder A, Pscheidt B, Haltrich D, Peterbauer C & Kittl R (2012) Simple and efficient expression of *Agaricus meleagris* pyranose dehydrogenase in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **94**, 695–704.
- 4 Hernández-Ortega A, Lucas F, Ferreira P, Medina M, Guallar V & Martínez AT (2012) Role of active site histidines in the two half-reactions of the aryl-alcohol oxidase catalytic cycle. *Biochemistry* **51**, 6595–6608.
- 5 Fan & Gadda G (2005) On the catalytic mechanism of choline oxidase. *J. Am. Chem. Soc.* **127**, 2067–2074.
- 6 Müller D (1928) Oxidation von Glukose mit Extrakten aus *Aspergillus niger*. *Biochem. Z.* **199**, 136–170.
- 7 Kujawa M, Ebner H, Leitner C, Hallberg BM, Prongjit M, Sucharitakul J, Ludwig R, Rudsander U, Peterbauer C, Chaiyen P, Haltrich D & Divne C (2006) Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J. Biol. Chem.* **281**, 35104–35115.
- 8 Cavener DR (1992) GMC oxidoreductases: A newly defined family of homologous proteins with diverse catalytic activities. *J. Mol. Biol.* **223**, 811–814.
- 9 Volc J, Kubátová E, Wood DA & Daniel G (1997) Pyranose 2-dehydrogenase, a novel sugar oxidoreductase from the basidiomycete fungus *Agaricus bisporus*. *Arch. Microbiol.* **167**, 119–125.
- 10 Volc J, Kubátová E, Daniel G, Sedmera P & Haltrich D (2001) Screening of basidiomycete fungi for the quinone-dependent sugar C-2/C-3 oxidoreductase, pyranose dehydrogenase, and properties of the enzyme from *Macrolepiota rhacodes*. *Arch. Microbiol.* **176**, 178–186.
- 11 Volc J, Sedmera P, Halada P, Daniel G, Přikrylová V & Haltrich D (2002) C-3 oxidation of non-reducing sugars by a fungal pyranose dehydrogenase: Spectral characterization. *J. Mol. Catal. - B Enzym.* **17**, 91–100.
- 12 Sedmera P, Halada P, Peterbauer C & Volc J (2004) A new enzyme catalysis: 3,4-dioxidation of some aryl β -D-glycopyranosides by fungal pyranose dehydrogenase. *Tetrahedron Lett.* **45**, 8677–8680.
- 13 Volc J, Sedmera P, Kujawa M, Halada P, Kubátová E & Haltrich D (2004) Conversion of lactose to β -D-galactopyranosyl-(1 \rightarrow 4)-D-arabino-hexos-2-ulose-(2-dehydrolactose) and lactobiono-1,5-lactone by fungal pyranose dehydrogenase. *J. Mol. Catal. B Enzym.* **30**, 177–184.
- 14 Sedmera P, Halada P, Kubátová E, Haltrich D, Přikrylová V & Volc J (2006) New biotransformations of some reducing sugars to the corresponding (di)dehydro(glycosyl)

- aldoses or aldonic acids using fungal pyranose dehydrogenase. *J. Mol. Catal. B Enzym.* **41**, 32–42.
- 15 Graf MMH, Bren U, Haltrich D & Oostenbrink C (2013) Molecular dynamics simulations give insight into D-glucose dioxidation at C2 and C3 by *Agaricus meleagris* pyranose dehydrogenase. *J. Comput. Aided Mol. Des.* **27**, 295–304.
 - 16 Pisanelli I, Kujawa M, Gschnitzer D, Spadiut O, Seiboth B & Peterbauer C (2010) Heterologous expression of an *Agaricus meleagris* pyranose dehydrogenase-encoding gene in *Aspergillus* spp. and characterization of the recombinant enzyme. *Appl. Microbiol. Biotechnol.* **86**, 599–606.
 - 17 Tan TC, Spadiut O, Wongnate T, Sucharitakul J, Krondorfer I, Sygmund C, Haltrich D, Chaiyen P, Peterbauer CK & Divne C (2013) The 1.6 Å crystal structure of pyranose dehydrogenase from *Agaricus meleagris* rationalizes substrate specificity and reveals a flavin intermediate. *PLoS ONE* **8**, e53567.
 - 18 Krondorfer I, Lipp K, Brugger D, Staudigl P, Sygmund C, Haltrich D & Peterbauer CK (2014) Engineering of pyranose dehydrogenase for increased oxygen reactivity. *PLoS ONE* **9**, e91145.
 - 19 Krondorfer I, Brugger D, Paukner R, Scheiblbrandner S, Pirker KF, Hofbauer S, Furtmüller PG, Obinger C, Haltrich D & Peterbauer CK (2014) *Agaricus meleagris* pyranose dehydrogenase: Influence of covalent FAD linkage on catalysis and stability. *Arch. Biochem. Biophys.* **558**, 111–119.
 - 20 Graf MMH, Lin Z, Bren U, Haltrich D, van Gunsteren WF & Oostenbrink C (2014) Pyranose dehydrogenase ligand promiscuity: A generalized approach to simulate monosaccharide solvation, binding, and product formation. *PLOS Comput. Biol.*, **10**, e1003995.
 - 21 Tokuriki N & Tawfik DS (2009) Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **19**, 596–604.
 - 22 Kelly SM, Jess TJ & Price NC (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **1751**, 119–139.
 - 23 Forneris F, Orru R, Bonivento D, Chiarelli LR & Mattevi A (2009) ThermoFAD, a ThermoFluor-adapted flavin *ad hoc* detection system for protein folding and ligand binding. *FEBS J.* **276**, 2833–2840.
 - 24 Wongnate T, Sucharitakul J & Chaiyen P (2011) Identification of a catalytic base for sugar oxidation in the pyranose 2-oxidase reaction. *ChemBioChem* **12**, 2577–2586.
 - 25 Chapman SK & Reid GA (1999) Flavoprotein Protocols. *Springer Science & Business Media*.
 - 26 Heuts DPHM, Scrutton NS, McIntire WS & Fraaije MW (2009) What's in a covalent bond? On the role and formation of covalently bound flavin cofactors. *FEBS J.* **276**, 3405–3427.
 - 27 Massey V (1991) Flavins and flavoproteins (Curti, B., Ronchi, S., & Zannetti, G., Eds.). *Walter de Gruyter & Co., Berlin*. pp 59-66.
 - 28 Mattevi A (2006) To be or not to be an oxidase: Challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* **31**, 276–283.
 - 29 McDonald CA, Fagan RL, Collard F, Monnier VM & Palfey BA (2011) Oxygen reactivity in flavoenzymes: Context matters. *J. Am. Chem. Soc.* **133**, 16809–16811.

- 30 Chaiyen P, Fraaije MW & Mattevi A (2012) The enigmatic reaction of flavins with oxygen. *Trends Biochem. Sci.* **37**, 373–380.
- 31 Braman J (ed.) (2002) *In vitro* mutagenesis protocols. *Humana Press, Totowa, New Jersey*.
- 32 Kujawa M, Volc J, Halada P, Sedmera P, Divne C, Sygmund C, Leitner C, Peterbauer C & Haltrich D (2007) Properties of pyranose dehydrogenase purified from the litter-degrading fungus *Agaricus xanthoderma*. *FEBS J.* **274**, 879–894.
- 33 Macheroux P, Massey V, Thiele DJ & Volokita M (1991) Expression of spinach glycolate oxidase in *Saccharomyces cerevisiae*: Purification and characterization. *Biochemistry* **30**, 4612–4619.
- 34 Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E, Springer BA, Lane P & Salemme FR (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screen.* **6**, 429–440.
- 35 Schmid N, Christ CD, Christen M, Eichenberger AP & van Gunsteren WF (2012) Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **183**, 890–903.
- 36 Oostenbrink C, Villa A, Mark AE & Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676.
- 37 Berendsen HJC, Postma JPM, Van Gunsteren WF & Hermans J (1981) Interaction models for water in relation to protein hydration. *Intermolecular Forces* **331**.
- 38 Amadei A, Chillemi G, Ceruso MA, Grottesi A & Di Nola A (2000) Molecular dynamics simulations with constrained roto-translational motions: Theoretical basis and statistical mechanical consistency. *J. Chem. Phys.* **112**, 9–23.
- 39 Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
- 40 Ryckaert J-P, Ciccotti G & Berendsen HJ. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.
- 41 Tironi IG, Sperb R, Smith PE & van Gunsteren WF (1995) A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459.
- 42 Heinz TN, van Gunsteren WF & Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **115**, 1125–1136.
- 43 Åqvist J, Medina C & Samuelsson JE (1994) A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **7**, 385–391.
- 44 Åqvist J, Luzhkov VB & Brandsdal BO (2002) Ligand binding affinities from MD simulations. *Acc. Chem. Res.* **35**, 358–365.
- 45 Stjernschantz E & Oostenbrink C (2010) Improved ligand-protein binding affinity predictions using multiple binding modes. *Biophys. J.* **98**, 2682–2691.

- 46 Fiehn O, Kopka J, Trethewey RN & Willmitzer L (2000) Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* **72**, 3573–3580.
- 47 Binh Chu D, Haberhauer-Troyer C, Mairinger T, Ortmayr K, Neubauer S, Koellensperger G & Han S (2014) Isotopologue analysis of sugar phosphates in yeast cell extracts by gas chromatography chemical ionization time of flight mass spectrometry. *J. Anal. Bioanal. Chem.*, in press.
- 48 Van Den Dool H & Dec. Kratz P (1963) A generalization of the retention index system including linear temperature programmed gas—liquid partition chromatography. *J. Chromatogr. A* **11**, 463–471.
- 49 Danneel H-J, Rossner E, Zeeck A & Giffhorn F (1993) Purification and characterization of a pyranose oxidase from the basidiomycete *Peniophora gigantea* and chemical analyses of its reaction products. *Eur. J. Biochem.* **214**, 795–802.
- 50 Sucharitakul J, Tongsook C, Pakotiprapha D, Berkel WJH van & Chaiyen P (2013) The reaction kinetics of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1 provide an understanding of the para-hydroxylation enzyme catalytic cycle. *J. Biol. Chem.* **288**, 35210–35221.
- 51 Chaiyen P, Brissette P, Ballou DP & Massey V (1997) Thermodynamics and reduction kinetics properties of 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase. *Biochemistry* **36**, 2612–2621.
- 52 Sucharitakul J, Wongnate T, Montersino S, van Berkel WJH & Chaiyen P (2012) Reduction kinetics of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1. *Biochemistry* **51**, 4309–4321.
- 53 Clark WM (1960) Oxidation-reduction potentials of organic systems (VJ Shiner Jr.). *Waverly Press, Inc., Baltimore*. pp 248-272.
- 54 Einarsdottir GH, Stankovich MT & Tu SC (1988) Studies of electron-transfer properties of salicylate hydroxylase from *Pseudomonas cepacia* and effects of salicylate and benzoate binding. *Biochemistry* **27**, 3277–3285.
- 55 Massey V, Komai H, Palmer G & Elion GB (1970) On the mechanism of inactivation of xanthine oxidase by Allopurinol and other pyrazolo[3,4-d]pyrimidines. *J. Biol. Chem.* **245**, 2837–2844.
- 56 Hutchison CA, Phillips S, Edgell MH, Gillam S, Jahnke P & Smith M (1978) Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**, 6551–6560.
- 57 Morrison KL & Weiss GA (2001) Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **5**, 302–307.
- 58 Tan T-C, Haltrich D & Divne C (2011) Regioselective control of β -D-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy. *J. Mol. Biol.* **409**, 588–600.
- 59 Higgins DR & Cregg JM (1998) *Pichia* protocols. *Springer Science & Business Media*.

- 60 Barrow CJ, Yasuda A, Kenny PTM & Zagorski MG (1992) Solution conformations and aggregational properties of synthetic amyloid β -peptides of Alzheimer's disease: Analysis of circular dichroism spectra. *J. Mol. Biol.* **225**, 1075–1093.
- 61 Tan T-C, Pitsawong W, Wongnate T, Spadiut O, Haltrich D, Chaiyen P & Divne C (2010) H-bonding and positive charge at the N(5)/O(4) locus are critical for covalent flavin attachment in *Trametes* pyranose 2-oxidase. *J. Mol. Biol.* **402**, 578–594.
- 62 Kenney WC, Edmondson DE & Seng RL (1976) Identification of the covalently bound flavin of thiamin dehydrogenase. *J. Biol. Chem.* **251**, 5386–5390.
- 63 Hassan-Abdallah A, Bruckner RC, Zhao G & Jorns MS (2005) Biosynthesis of covalently bound flavin: Isolation and *in vitro* flavinylation of monomeric sarcosine oxidase apoprotein. *Biochemistry* **44**, 6452–6462.
- 64 Schleicher E, Hitomi K, Kay CWM, Getzoff ED, Todo T & Weber S (2007) Electron nuclear double resonance differentiates complementary roles for active site histidines in (6-4) photolyase. *J. Biol. Chem.* **282**, 4738–4747.
- 65 Hecht HJ, Kalisz HM, Hendle J, Schmid RD & Schomburg D (1993) Crystal structure of glucose oxidase from *Aspergillus niger* refined at 2.3 Å resolution. *J. Mol. Biol.* **229**, 153–172.
- 66 Bren U & Oostenbrink C (2012) Cytochrome P450 3A4 inhibition by ketoconazole: Tackling the problem of ligand cooperativity using molecular dynamics simulations and free-energy calculations. *J. Chem. Inf. Model.* **52**, 1573–1582.
- 67 Vasanthanathan P, Olsen L, Jørgensen FS, Vermeulen NPE & Oostenbrink C (2010) Computational prediction of binding affinity for CYP1A2-ligand complexes using empirical free energy calculations. *Drug Metab. Dispos.* **38**, 1347–1354.
- 68 Shirts MR, Mobley DL & Brown SP (2010) Free energy calculations in structure-based drug design. *Drug Des. Struct.-Ligand-Based Approaches, Cambridge*. pp 61–86.
- 69 De Ruiter A & Oostenbrink C (2012) Efficient and accurate free energy calculations on trypsin inhibitors. *J. Chem. Theory Comput.* **8**, 3686–3695.
- 70 Ruiz-Matute AI, Hernández-Hernández O, Rodríguez-Sánchez S, Sanz ML & Martínez-Castro I (2011) Derivatization of carbohydrates for GC and GC–MS analyses. *J. Chromatogr. B* **879**, 1226–1240.
- 71 Little JL (1999) Artifacts in trimethylsilyl derivatization reactions and ways to avoid them. *J. Chromatogr. A* **844**, 1–22.
- 72 Volc J, Sedmera P, Halada P, Příkrylová V & Daniel G (1998) C-2 and C-3 oxidation of D-Glc, and C-2 oxidation of D-Gal by pyranose dehydrogenase from *Agaricus bisporus*. *Carbohydr. Res.* **310**, 151–156.
- 73 Wilke T, Schneider M & Kleinermanns K (2013) 1,4-hydroquinone is a hydrogen reservoir for fuel cells and recyclable *via* photocatalytic water splitting. *Open J. Phys. Chem.* **03**, 97–102.
- 74 Harcourt AV (1867) On the observation of the course of chemical change. *J. Chem. Soc.* **20**, 460–495.
- 75 Wolfenden R & Snider MJ (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* **34**, 938–945.

-
- 76 Fraaije MW, Heuvel RHH van den, Berkel WJH van & Mattevi A (1999) Covalent flavinylation is essential for efficient redox catalysis in vanillyl-alcohol oxidase. *J. Biol. Chem.* **274**, 35514–35520.
- 77 Edmondson DE & Ghisla S (1999) Flavins and Flavoproteins (Ghisla, S., Kroneck, P., Macheroux, P., and Sund, H. Eds.). *Rudolf Weber, Berlin*. pp 71–76.
- 78 Tasca F, Gorton L, Kujawa M, Patel I, Harreither W, Peterbauer CK, Ludwig R & Nöll G (2010) Increasing the coulombic efficiency of glucose biofuel cell anodes by combination of redox enzymes. *Biosens. Bioelectron.* **25**, 1710–1716.
- 79 Gadda G (2012) Oxygen activation in flavoprotein oxidases: The importance of being positive. *Biochemistry* **51**, 2662–2669.
- 80 Baron R, Riley C, Chenprakhon P, Thotsaporn K, Winter RT, Alfieri A, Forneris F, Van Berkel WJH, Chaiyen P, Fraaije MW, Mattevi A & McCammon JA (2009) Multiple pathways guide oxygen diffusion into flavoenzyme active sites. *Proc. Natl. Acad. Sci.* **106**, 10603–10608.
- 81 Hernandez-Ortega A, Lucas F, Ferreira P, Medina M, Guallar V & Martinez AT (2011) Modulating O₂ reactivity in a fungal flavoenzyme: Involvement of aryl-alcohol oxidase Phe-501 contiguous to catalytic histidine. *J. Biol. Chem.*, jbc.M111.282467.
- 82 Finnegan S, Agniswamy J, Weber IT & Gadda G (2010) Role of valine 464 in the flavin oxidation reaction catalyzed by choline oxidase. *Biochemistry* **49**, 2952–2961.

5.6 Supplementary material

Table S1: Nucleotide sequences of the primers used in this study. The primers were necessary for removing the c-myc epitope of the pPICZ B vector, adding the NotI and XbaI restriction sites to the 3'- and 5'-ends of the ampdh gene, respectively, and for site-directed mutagenesis. Restriction and mutagenesis sites are highlighted in bold, underlined letters.

Name	Sequence (5'-3')
pPICZB-6His-fw	CATCATCATCATCATCATTGAGTTTGTAGCCTTAGACATG
pPICZB-6His-XbaI-rv	ATGATGATGATGATGTCTAGAAAGCTGGCGGCCGCCGCGGC
AmpDH-NotI-fw	AAAGCGGCCCGCATGCTGCCTCGAGTGACCAAGTTG
AmpDH-XbaI-rv	TTTTCTAGAGTTATAACTCTTTGCTATCAACGC
H103A-fw	GGTGGCTGCAGTACTGCTAATGGAATGGTGTACACC
H103A-rv	AGTACTGCAGCCACCCAGGATCTTTGC
Q392A-fw	CCACATATTGAGTTCGCTTTTGCACAAATCACCCC
Q392A-rv	GAACCAATATGTGGTGAATTCCTTGCC
Y510A-fw	CITCAACATTCCTCAGCTGTGCATGGTGTGGGAACG
Y510A-rv	TGAGAATGTTGAAGATCGCAGG
V511F-fw	CAACATTCTCATACTTCCATGGTGTGGGAACGTTG
V511W-fw	CAACATTCTCATACTGGCATGGTGTGGGAACGTTG
V511-rv	GTATGAGAATGTTGAAGATCGCAGG
H512A-fw	ATTCTCATAACGTGGCTGGTGTGGGAACGTTGTCG
H512A-rv	CACGTATGAGAATGTTGAAGATCG
H556A-fw	GCTCCGGCCCGCAGCTACTCAACTACCTGTTTACGC
H556A-rv	AGTAGCTGCGGCCGGAGCATGC
H556N-fw	GCTCCGGCCCGCAAACACTCAACTACCTGTTTACGC
H556N-rv	AGTGTTTGCAGGCCGGAGCATGC

Table S2: Retention indices and mass accuracy of derivatization products of D-glucose (GLC), mono- and diketo-GLC. More details on sample preparations and measurements are listed in the materials and methods (section 5.2.10).

Analytes	Chemical formula of quasi-molecular ion	Retention index	Theoretical mass [Da]	Measured mass [Da]	Mass accuracy [ppm]
2-keto-GLC	[C ₂₂ H ₅₃ N ₂ O ₆ Si ₄] ⁺	2162	553.2975	553.2981	1.0
3-keto-GLC	[C ₂₂ H ₅₃ N ₂ O ₆ Si ₄] ⁺	2192	553.2897	553.3050	9.5
2,3-diketo-GLC	[C ₂₁ H ₄₈ N ₃ O ₆ Si ₃] ⁺	2157	522.2845	522.2850	0.9
GLC	[C ₂₃ H ₅₈ O ₆ Si ₅] ⁺	2188	584.3105	584.3092	2.2

Table S3: Parameter sets for simulating stopped-flow traces with the program Berkeley Madonna (version 8.3.14) according to material and methods (section 5.2.11). The reductive half-reaction recorded for different electron donors (sugars) was simulated in agreement with kinetic model 1_{red}. Abbreviations for utilized sugars: D-glucose (GLC), D-galactose (GAL), L-arabinose (ARA), and methyl- α -D-glucopyranoside (MGP). The oxidative half-reaction for the electron acceptor 1,4-benzoquinone (BQ) was simulated according to kinetic model 1_{ox}. Simulated traces of the reductive- and oxidative half-reactions for AmPDH are shown as red, dashed lines in **Figure S2** or **Figure S3**, respectively, as well as in **Figure S4** for variant H556A.

Variant	Substrate	k_1 [M ⁻¹ s ⁻¹]	k_2 [M ⁻¹ s ⁻¹]	ϵ_A [M ⁻¹ cm ⁻¹]	ϵ_B [M ⁻¹ cm ⁻¹]	ϵ_C [M ⁻¹ cm ⁻¹]
AmPDH	GLC	6.9×10^4	5.4	9.0×10^3	1.1×10^3	0.7×10^3
	GAL	5.2×10^4	4.0	9.3×10^3	1.0×10^3	0.8×10^3
	ARA	7.8×10^4	5.0	9.5×10^3	1.5×10^3	1.5×10^3
	MGP	6.5×10^4	1.0	9.3×10^3	1.5×10^3	1.3×10^3
	BQ	8.0×10^3	0.05	1.7×10^3	1.3×10^4	9.2×10^3
H556A	GLC	1.0×10^2	8.0×10^1	1.3×10^4	9.8×10^3	1.1×10^3
	BQ	3.6×10^4	1.6×10^3	1.2×10^3	2.0×10^4	1.1×10^4

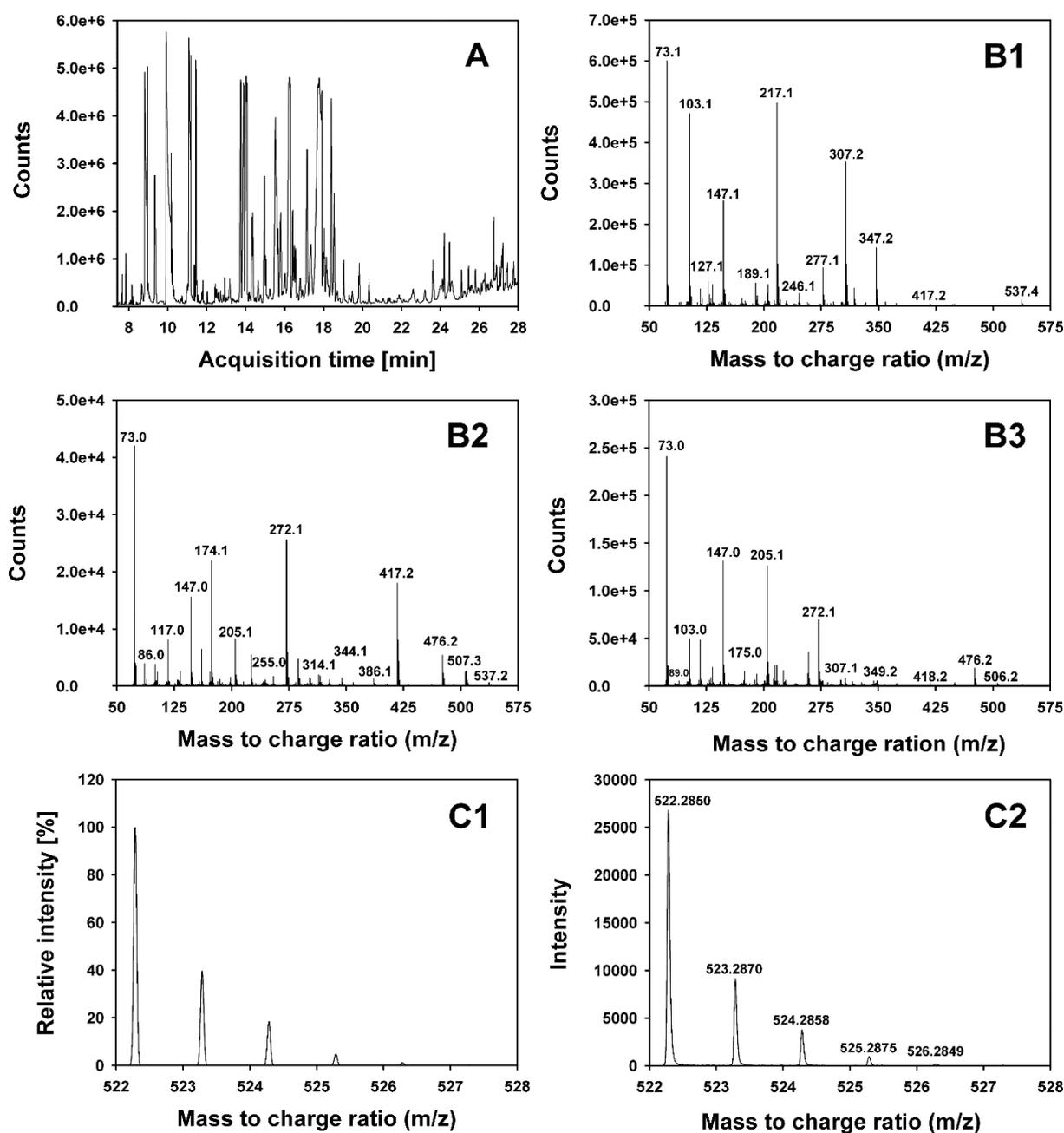


Figure S1: GC-MS measurements. (A) Total ion chromatogram of the AmPDH reaction product after 120 min on the HP1MS column with electron ionization mass spectrometry detection. Electron ionization mass spectra of two-step derivatizations of 2-keto-D-glucose (B1), 3-keto-D-glucose (B2) and 2,3-diketo-D-glucose (B3). Isotopologue fraction of the protonated molecular ion of ethoxymation-trimethylsilylation derivatization products of 2,3-diketo-D-glucose: theoretical distribution (A) and experimental distribution (B). More details on sample preparations and measurements are listed in the materials and methods (section 5.2.10).

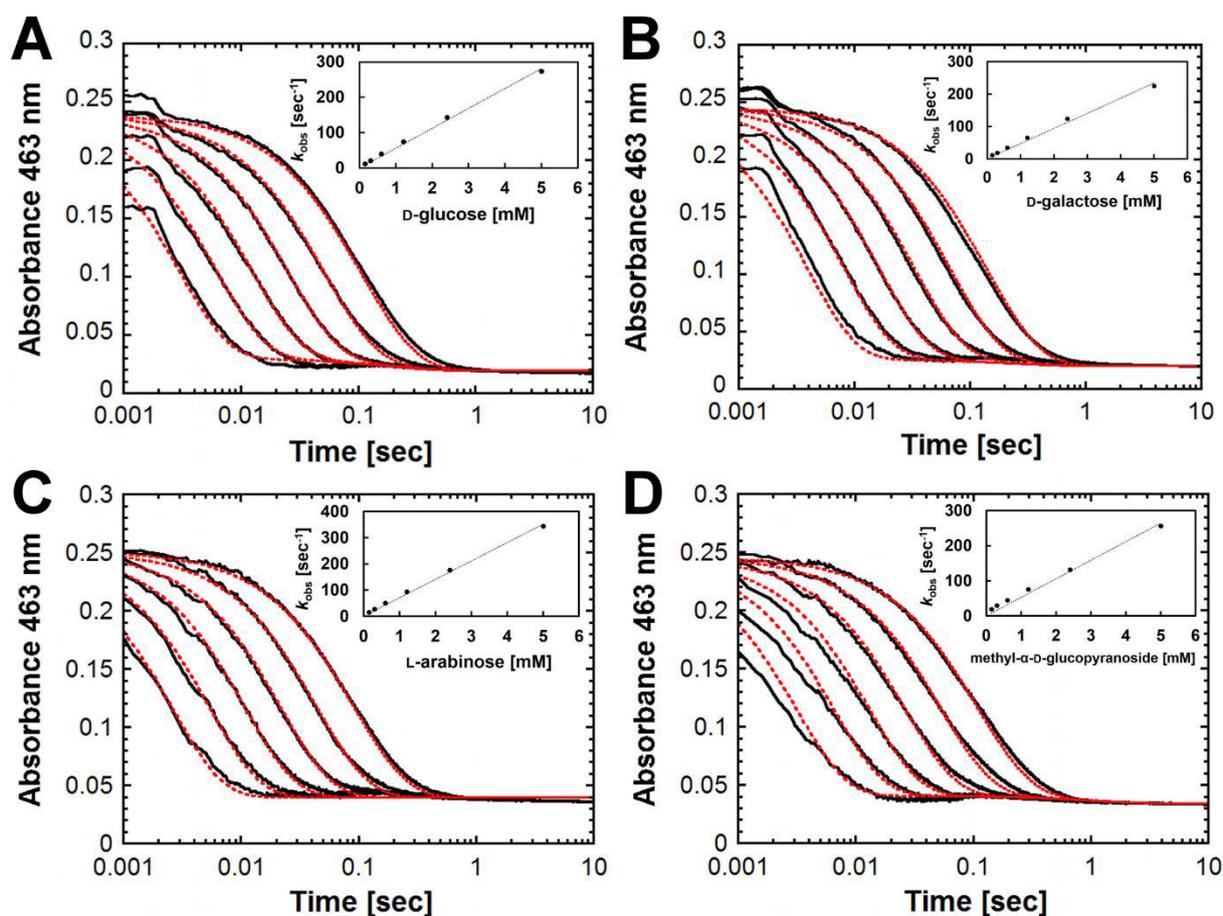


Figure S2: Reductive half-reaction of AmPDH followed by stopped-flow spectroscopy after mixing with sugar substrates (A) D-glucose, (B) D-galactose, (C) L-arabinose, and (D) methyl- α -D-glucopyranoside. A solution of fully oxidized enzyme ($30 \mu\text{M}$) was mixed with buffer containing various sugar concentrations: 0.15, 0.3, 0.6, 1.2, 2.4, and 5 mM in 50 mM potassium phosphate buffer (pH 7.0) at 4°C . All given concentrations are after mixing. The reaction was monitored at 463 nm with a stopped-flow spectrophotometer. The kinetic traces from right to left correspond to increasing sugar concentrations. Color code for traces: black (experiment); red (simulation with the program Berkeley Madonna (version 8.3.14) according to the kinetic model 1_{red} as mentioned in material and methods (section 5.2.11) with the parameter sets listed in **Table S3**). The inset shows the plot of the pseudo-first-order rate constants versus the concentrations of the respective sugar.

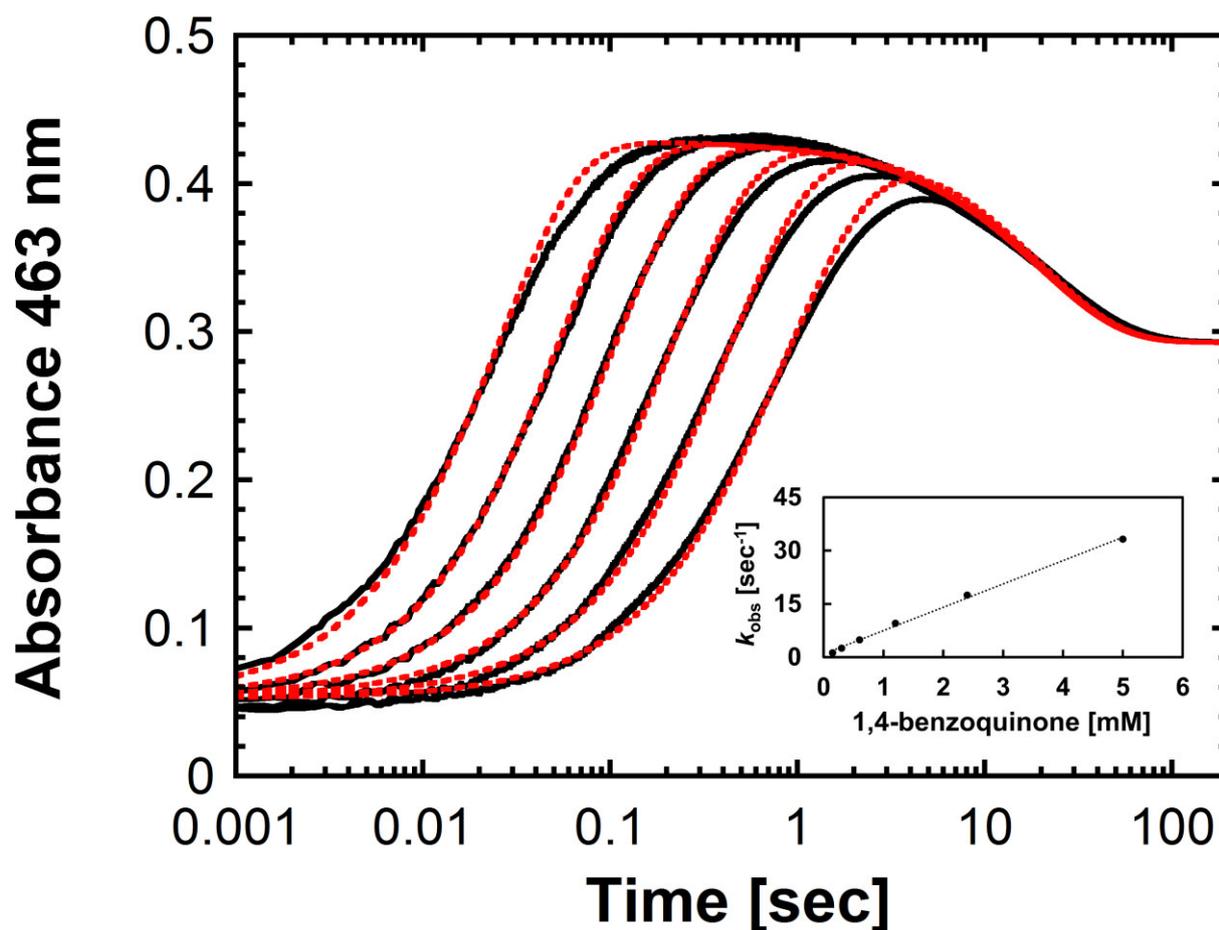


Figure S3: Oxidative half-reaction of AmPDH followed by stopped-flow spectroscopy after mixing with 1,4-benzoquinone (BQ). A solution of fully reduced enzyme (30 μM) was mixed with buffer containing various BQ concentrations: 0.15, 0.3, 0.6, 1.2, 2.4, and 5 mM in 50 mM potassium phosphate buffer (pH 7.0) at 4°C under anaerobic conditions. All given concentrations are after mixing. The reaction was monitored at 463 nm with a stopped-flow spectrophotometer. The kinetic traces from right to left correspond to increasing BQ concentrations. Color code for traces: black (experiment); red (simulation with the program Berkeley Madonna (version 8.3.14) according to the kinetic model 1_{ox} as mentioned in material and methods (section 5.2.11) with the parameters listed in **Table S3**. Simulations were conducted until 10 sec of reaction time, as the last phase was not considered relevant for the oxidative half-reaction of AmPDH. The inset shows the plot of the pseudo-first-order rate constants versus BQ concentrations.

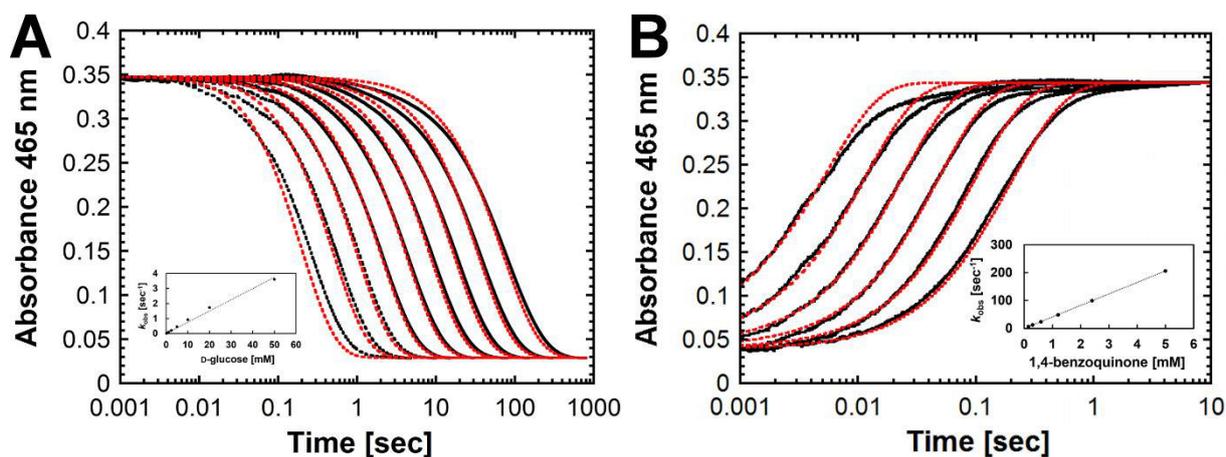


Figure S4: Reductive and oxidative half-reaction of 30 μM H556A followed by stopped-flow spectroscopy after mixing with (A) D-glucose (GLC) or (B) 1,4-benzoquinone (BQ) in 50 mM potassium phosphate buffer (pH 7.0) at 4°C under anaerobic conditions. The reaction was monitored at 465 nm with a stopped-flow spectrophotometer. All given concentrations are after mixing. Color code for traces: black (experiment); red (simulation with the program Berkeley Madonna (version 8.3.14) according to (A) kinetic model 1_{red} or (B) kinetic model 1_{ox} as mentioned in material and methods (section 5.2.11) with the parameter sets listed in **Table S3**. The inset shows the plot of the pseudo-first-order rate constants versus varying (A) GLC or (B) BQ concentrations. (A) A solution of fully oxidized H556A was mixed with buffer containing various GLC concentrations: 0.15, 0.3, 0.6, 1.2, 2.4, 5, 10, 20, and 50 mM. The kinetic traces from right to left correspond to increasing GLC concentrations. (B) A solution of fully reduced H556A was mixed with buffer containing various BQ concentrations: 0.15, 0.3, 0.6, 1.2, 2.4, and 5 mM. The kinetic traces from right to left correspond to increasing BQ concentrations.

CHAPTER 6

Conclusion and Outlook

"He did not arrive at this conclusion by the decent process of quiet, logical deduction, nor yet by the blinding flash of glorious intuition, but by the shoddy, untidy process halfway between the two by which one usually gets to know things."

Margery Allingham

Author

Michael Graf

6.1 Conclusion of the thesis

In the course of this PhD thesis, experimental and computational approaches were combined to comprehensively investigate the active site of *Am*PDH. Consequently, the conclusions could be drawn from either an experimental or a computational point of view. Since the aim of this PhD thesis was to probe the active site of *Am*PDH – irrespective of the applied method – the final conclusion will be presented in an integrated manner.

This is especially feasible for **Chapter 3** and **Chapter 5**. In the former chapter, MD simulations suggested mutagenesis experiments, which were performed in the course of **Chapter 5**. Therein, a multitude of different experimental techniques was applied to obtain biochemical and biophysical insights of active site residues in *Am*PDH. Whenever possible, these findings were compared to MD simulations. A single catalytic base for D-glucose (GLC) oxidation and the importance of other active site residues for GLC binding could be elucidated and corroborated. Interestingly, the GLC oxidation sites of all variants investigated in this PhD thesis were the same compared to the wild type. This outcome was unexpected, since MD simulations suggested that single amino acid exchanges could potentially affect the ratio of C2 and C3 oxidation. This indicates that several residues are involved in GLC binding and alterations of one of them can be compensated by the others. The flexibility in substrate binding could be rationalized by MD simulations in **Chapter 3**, where a diverse H-bonding pattern between different active site residues in wild type *Am*PDH and GLC was observed. Moreover, van der Waals interactions were identified as the main driving force for substrate binding, which was observed previously for promiscuous enzymes. The substrate promiscuity of *Am*PDH was subject of another study employing extensive MD simulations, which are presented in **Chapter 4**. Therein, experimentally detected sugar oxidation products could be reproduced and new ones predicted, suggesting new experiments. Stopped-flow measurements described in **Chapter 5** indicated that *Am*PDH equally prefers C2 or C3 oxidized substrates and that the reductive half-reaction is the main determinant for decreased catalytic efficiency of variant H556A. Interestingly, most studied variants had a higher oxygen reactivity compared to wild type *Am*PDH, which is still low when compared to ‘true’ oxidases, though, indicating that (i) the wild type was efficiently optimized in the course of evolution to be a dehydrogenase and that (ii) the chemical environment around the isoalloxazine is indeed important for modulating its oxygen reactivity. Finally, mutational studies demonstrated that covalent flavinylation increases the redox potential of FAD in *Am*PDH, which is in agreement with similar studies on other flavoenzymes.

In this PhD thesis, *Am*PDH was extensively characterized by experimental and computational means. As pointed out in the introduction of this PhD thesis, such an integrated approach is in line with recent studies in the area of (computational) protein engineering to gain detailed fundamental insights of the studied enzyme and engineer it towards desired functions.

6.2 Outlook

The aim of this PhD thesis was to characterize the active site of *Am*PDH by biochemical, biophysical, and computational means. The ultimate goal was to obtain detailed fundamental biochemical insights to provide better approaches for the engineering of *Am*PDH towards its desired applications.

In this context, an interesting aspect will be to generate additional active site variants, which are based on the data obtained in the course of this PhD thesis. Thereby, substrate oxidation sites could possibly be altered or new ones introduced, which would certainly be beneficial for applications in organic synthesis and bioelectrochemistry. More detailed investigations on the increased oxygen reactivities for variants H103A and H556A will be interesting to further rationalize this – still elusive – property of flavoenzymes. Generating the double mutant H103A/H556A might yield *Am*PDH variants with even higher oxygen reactivities compared to the single mutants H103A and H556A, which could lead to a better understanding of the oxygen reactivity in *Am*PDH.

As pointed out in the introduction, QM/MM calculations would be interesting to perform in order to obtain more insights into the hydride-transfer from the sugar substrates to the flavin N5-atom. Such calculations would also allow to determine the relative energy profiles for the reaction of *Am*PDH with O₂, which could potentially provide a better understanding of the oxygen reactivity of *Am*PDH.

From a computational point of view, faster hardware will allow for increased simulation times of more complex systems. This will eventually contribute to make a combined experimental and computational approach even more feasible for similar studies on other enzymes as presented in this PhD thesis.

List of Publications

Chapter 3

Graf MMH*, Bren U*, Haltrich D & Oostenbrink C (2013) Molecular dynamics simulations give insight into D-glucose dioxidation at C2 and C3 by *Agaricus meleagris* pyranose dehydrogenase. *J. Comput. Aided Mol. Des.*, **27**, 295–304. *Both authors contributed equally to this work.

Chapter 4

Graf MMH, Lin Z, Bren U, Haltrich D, van Gunsteren WF & Oostenbrink C (2014) Pyranose dehydrogenase ligand promiscuity: A generalized approach to simulate monosaccharide solvation, binding, and product formation. *PLOS Comput. Biol.*, **10**, e1003995.

Chapter 5

Graf MMH, Sucharitakul J, Bren U, Binh CD, Koellensperger G, Hann S, Furtmüller PG, Obinger C, Oostenbrink C, Chaiyen P, Haltrich D (2015) Active site residues in *Agaricus meleagris* pyranose dehydrogenase: Structure–dynamics–function relationships unraveled by biochemical, biophysical, and computational means. In preparation.

Curriculum Vitae

Michael Graf, Dipl.-Ing.

Education

- 12/10 – 01/15** PhD studies in Biochemistry at the Food Biotechnology Laboratory at the Universität für Bodenkultur (BOKU), Vienna, Austria
Thesis title: Probing the active site of pyranose dehydrogenase by experimental and computational means (Supervisor: Prof. Dietmar Haltrich)
- 12/13 – 02/14** Research stay at Mahidol University in Bangkok, Thailand
- 03/13 – 05/13** Research stay at ETH in Zurich, Switzerland
- 10/07 – 08/10** Master studies in Biotechnology at BOKU, Vienna
Thesis title: Platform technology for the production of therapeutic antibody fragments in Escherichia coli (Supervisors: Prof. Diethard Mattanovich in cooperation with Dr. Hubertus Hobenblum from Boehringer Ingelheim, Vienna)
- 03/10 – 06/10** Study visit at the University of Bologna, Italy
- 08/09 – 12/09** Study visit at Cornell University, Ithaca, NY, USA
- 10/03 – 09/07** Bachelor studies in Food Science and Biotechnology at BOKU, Vienna
- 09/94 – 06/02** Secondary school, BRG/BORG Dornbirn Schoren, Austria
- 08/99 – 06/00** Study visit at Clay County High School, Ashland, AL, USA

Scientific Conferences

- 09/2014** 6th ÖGMBT Annual Meeting, Vienna, Austria; Poster
- 09/2014** 24th Intl. BIOMOS Symposium on Biomolecular Simulation, Ausserberg, Switzerland; Talk
- 07/2014** 18th Intl. Symposium on Flavins and Flavoproteins, Phechaburi, Thailand; Poster
- 07/2014** OxiZymes, Vienna, Austria; Talk
- 09/2012** OxiZymes, Marseille, France; Poster
- 06/2012** Trends in Enzymology, Göttingen, Germany; Poster
- 07/2011** 17th Intl. Symposium on Flavins and Flavoproteins, Berkeley, CA, USA; Poster
- 02/2011** 6th Conference on Recombinant Protein Production, Vienna, Austria

Acknowledgements

I would like to acknowledge the financial support from universities and funding agencies, in particular the Austrian Science Fund (FWF), which funded my PhD in the course of the doctoral program Biomolecular Technology of Proteins (BioToP), the University of Natural Resources and Life Sciences, Vienna (BOKU), the Swiss Federal Institute of Technology (ETH) in Zurich, and the Mahidol University in Bangkok. Thanks to following people for their support:

Supervisor

Univ. Prof. Dietmar Haltrich

PhD committee members

Univ. Prof. Pimchai Chaiyen
Univ. Prof. Christina Divne
Univ. Prof. Christian Obinger

Examiners

Univ. Prof. Christian Obinger
Assoc. Prof. Bojan Zagrovic

LBT group members

Priv.-Doz. Clemens Peterbauer
Dr. Roland Ludwig
Dr. Roman Kittl
Dr. Christoph Sygmund
Dr. Ha Nguyen Thu
Dr. Su Ma
Dr. Herbert Michlmayr
Dr. Sheryl Arreola
Dr. Anh Nguyen Hoang
Dr. Montira Intanon
Dr. Apinun Kanpiengjai
Dr. Vanja Kaswurm
Dr. Tien-Thanh Nguyen
Dr. Iqbal Sanaullah
Dr. Leonard Stoica
Dagmar Brugger
Alfons Felice
Barbara Geiger

Christoph Gonaus
Hung
Anna Kowalik
Daniel Kracher
Iris Krondorfer
Minh Nguyen
Regina Paukner
Aleksandra Pinczewska
Marita Preims
Stefan Scheiblbrandner
Petra Staudigl
Kawah Zahma

Supervised students

Kristina Fuchs
Michael Kitzmüller
Stefan Klingohr
Claudia Lechner
Kathrin Schrotta
Benjamin Sevcnikar
Sandra Weber

Group supporting members

Dagmar Fort
Katharina Grünböck
Victoria Hell
Cindy Lorenz
Helena Szabo

MMS group members

Univ. Prof. Chris Oostenbrink

Assoc. Prof. Urban Bren

Dr. Antonio Garate
Dr. Melanie Grandits
Dr. Balder Lai
Dr. Gabor Nagy
Dr. Maria Pechlaner
Dr. Maria Reif
Dr. Anita de Ruiter
Dr. Axel Sündermann
Wolfgang Buchta
Matthias Diem
Christian Margreitter
Manuela Maurer
Martina Setz
Sonja Wit

BioToP and BOKU members

Univ. Prof. Paul Furtmüller
Assoc. Prof. Stephan Hann
Univ. Prof. Gunda Köllensperger
Dinh Binh Chu

External collaborators

Univ. Prof. Pimchai Chaiyen (TH)
Univ. Prof. Wilfred van Gunsteren (CH)
Assoc. Prof. Jeerus Sucharitakul (TH)
Dr. Monika Laner (CH)
Dr. Tien-Chye Tan (SE)
Dr. Jindrich Volc (CZ)
Dr. Bjørge Westereng (NO)
Dr. Lin Zhixiong (CH)

Special thanks are extended to Dietmar, who gave me the opportunity to do the PhD in his group, for supervision, for the connections to Mahidol University, and for letting my research develop towards my interests without any constraints. Moreover, I owe thanks to Iris for introducing me to the wonders of PDH, to Roman and Christoph (Sygmund) for showing me the workings of the MBR- and Applikon fermenters and especially Christoph (Sygmund) and Herbert for advising me in molecular biology. Daniel for support in handling the ÄKTAs and other lab-equipment and Dagmar (Brugger), especially for clarifications regarding the photometers. Thanks to Sheryl and Mon for help in TLC as well as Cindy and Victoria for HPLC and ordering chemicals. Very special

thanks to the bachelor- and master students I could supervise during my PhD – I probably learned as much from them as they did (hopefully) from me. Miss Szabo, Dagmar (Fort) and Katharina for being great secretaries. Vanja, Fon, Kawah, and Thanh were a great support and cheered me up with their bright attitude – especially in the beginning of my PhD, when more experiments went wrong than right. Thanks to Dagmar (Brugger), Herbert, Christoph (Gonaus), and volatile others for fun and interesting conversations over lunch in the ‘Krone-Mensa’. Hung and Minh (aka ‘sister’) for being great office colleagues. Gratitude goes to the all LBT lab members for occasional celebrations and after-work beers.

Thanks to Christina and TC for providing a preliminary X-ray structure of PDH and advice at the beginning of my PhD about which positions to select for mutagenesis studies. Christian (Obinger) and Paul I would like to thank for help with biochemical and biophysical questions as well as Stephan (Hann), Gunda, and Binh for extensive GC-MS measurements. For further assistance concerning the detection of D-glucose oxidation products, I also owe thanks to Jindra and Bjorge.

Pimchai and all PC lab members from Mahidol University showed me how far hospitality actually can go and managed to make me feel at home in Bangkok – despite the huge protests on the streets (‘Shutdown Bangkok’). Moreover, they introduced me to basics in Buddhism and Thai history (special thanks to Pom and Mamuang) and many other aspects such as genuine, spicy Thai food. Additionally, I could learn a great deal about redox potential measurements and pre-steady-state kinetics of flavoenzymes from true experts in that field, especially Pimchai and Pom. Thank you all for letting me have such an extraordinary experience in beautiful Thailand!

These acknowledgements cannot be complete without giving special thanks to Chris and Urban, who introduced me to MD simulations and greatly supported me all along the way. You both gave my PhD thesis an unexpected, yet exciting kink, which I very much appreciate. Thanks to all present and former MMS group members as well as Daniela for giving me a cozy, second home at BOKU. Thanks to Chris for the connections to ETH Zurich and to Wilfred for accepting me as a guest researcher there. The period in Zurich was very exciting, both personally (thanks Indrani, Kingshuk, Sia, and the whole IGC group) and scientifically (thanks Wilfred, Alan, Zhixiong, and Monika).

Last but not least I want to thank my family, Becky & Jerry, Gudrun’s family, and my friends for their continued support and for being there! In the end, I extend my gratitude to Gudrun, who supported me all along the way and was always there for me – no matter what. Gudrun, it is good to have someone like you in my life!