

Bryn Mawr College

## Scholarship, Research, and Creative Work at Bryn Mawr College

---

Bryn Mawr College Dissertations and Theses

---

2022

### Understanding and Harnessing the Structural Features that Govern HMG-CoA Reductase Activity

Edwin Ragwan  
*Bryn Mawr College*

Follow this and additional works at: <https://repository.brynmawr.edu/dissertations>

 Part of the [Chemistry Commons](#)

---

#### Custom Citation

Ragwan, E. 2022. "Understanding and Harnessing the Structural Features that Govern HMG-CoA Reductase Activity." PhD Diss., Bryn Mawr College.

This paper is posted at Scholarship, Research, and Creative Work at Bryn Mawr College.  
<https://repository.brynmawr.edu/dissertations/236>

For more information, please contact [repository@brynmawr.edu](mailto:repository@brynmawr.edu).

# **Understanding and Harnessing the Structural Features that Govern HMG-CoA Reductase Activity**

by

Edwin Ragwan

2022

Submitted to the Faculty of Bryn Mawr College  
in partial fulfillment of the requirements for  
the degree of Doctor of Philosophy  
in the Department of Chemistry

Doctoral Committee:

Yan Kung, Advisor

Susan A. White

Sharon Burgmayer

Patrick R. Melvin

## Abstract

The enzyme 3-hydroxy-3-methylglutaryl coenzyme-A reductase (HMGR), which catalyzes the reduction of HMG-CoA to mevalonate using two equivalents of the cofactor NAD(P)H, is part of the mevalonate pathway, found in all kingdoms of life. This pathway is at the heart of natural product biosynthesis constituting one of the obligate routes to producing the building blocks for isoprenoids, which represent the largest and most diverse class of natural products. Natural products and its derivatives continue to provide important compounds in the fields of drug development, biomedical engineering, and commercially-driven products. Therefore, HMGR, which performs the rate-limiting step of the mevalonate pathway garners significant attention in these fields. Understanding of the mechanism of HMGR remains incomplete, with gaps pertaining to the role of a putative C-terminal flap domain (CTD) responsible for helping to modulate the positioning of active site residues as well as the flux of substrates, intermediates, and products during the reaction. Moreover, investigation into structural determinants of cofactor specificity, which contribute towards varied cofactor preferences observed among HMGR homologs, point to a cofactor helix as being crucial, and warrants further analysis.

Here, class II HMGRs from *Delftia acidovorans* (DaHMGR) and *Streptococcus pneumoniae* (SpHMGR) are used as case-studies in elucidating biophysical, kinetic, and X-ray crystallographic features that aid in our understanding of the three-dimensional characteristics of class II HMGR including its catalytic and oligomeric states in solution, the role of the CTD in the mechanism and features governing cofactor specificity. We provide the first crystal structures of

DaHMGR alongside complementary kinetic studies showing it to be an NAD-preferring HMGR. We observe that NAD-preferring HMGRs are able to form hexamers whereas NADP-preferring HMGRs predominate as dimers. Structures obtained with the CTD in novel locations expands our knowledge on this domain. We observe a flipped conformation of the CTD that appears to capture an intermediate state. The importance of the cofactor helix is manifest through protein engineering efforts that switch specificity by switching this motif between HMGRs. This work expands on the current class II HMGR paradigm as it relates to structural and functional dynamics, providing greater insight into this biologically important enzyme.

“There is nothing like returning to a place that remains unchanged to find the ways in which you yourself have altered.”

Nelson Mandela

## Acknowledgements

I am fortunate to have had Dr. Yan Kung as an adviser, who has extended to me layers of mentoring ranging from encouraging and supportive words to the technical skills involved in the lab, to scientific pragmatism and curiosity. I am grateful to you for allowing me to grow in an inviting space, which propelled me towards my own scientific endeavors as well as in my roles within the lab. Thank you for giving me both the guidance and freedom in this regard. I am sincerely grateful for all you have done for and behalf of me.

I am also fortunate to have had Dr. Sharon Burgmayer, Dr. Greg Davis, Dr. Patrick Melvin, and Dr. Susan White on my committee. Thank you for shaping my academic career during this season, for your time, and for the nurturing sentiments expressed at various milestones. I am grateful for members of the Bryn Mawr Community, the Chemistry Department, and the staff and faculty, for their direct contributions towards my success. It is an understatement to say that without you all this could not have been done.

Throughout my time as a graduate student and teaching assistant, I have been privileged to collaborate with many colleagues, undergraduate, post-baccalaureate, and graduate students. My fellow graduate students have always been there to remind me that we are all in this together. The undergraduate students in the Kung Lab, have been instrumental in developing me as a graduate student. I am sincerely appreciative of the trust, and mutual respect, and collegiate atmosphere that was present in the lab, and it was especially meaningful given the diverse backgrounds that were represented. I always felt valued and affirmed in the lab.

I am deeply grateful to my parents, who from an early age, instilled so many good priorities and values, that eventually informed my worldview in and out of the lab. Thank you for doing your best in providing the opportunities that I benefited from, for your encouragement that came in so many forms, and for modeling for me what a steward of such pathways should look like. Thank you to my wife's parents, too; I always knew I could count on your prayers and support. Likewise, my siblings and their respective family have extended a measure of support that not only came by way of care and concern, but also by your consistent presence. Knowing that you were around encouraged me in profound ways, and you should know how much this meant to me.

My children, Micah, and Anna have motivated me in ways it is hard to express in words. Your ever-present love for me is humbling, and it is a great honor to have you in my lives. Coming home to you at the end of the day was a highlight that never got old. Finally, I am most fortunate to have been supported and cared for by my wife, Jehnell. Thank you for doing small things like coming to the lab on weekends when I needed to set up experiments, and for bigger things like caring for our kids when I was away. Your actions, and your words, were the most meaningful gifts to me. Your steadfast presence, along with my faith in God, through various experiences was invaluable. Your optimism and the confidence you gave me, offered me hope during times of self-doubt. You are a gem, and I will always be deeply grateful for you.

## Table of Contents

<b>Abstract .....</b>	<b>i</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>vi</b>
<b>Chapter I: HMG-CoA reductase, a key enzyme in natural product synthesis.....</b>	<b>1</b>
<b>1.1. Natural Products .....</b>	<b>1</b>
<b>1.2. Isoprenoid Biosynthesis .....</b>	<b>2</b>
1.2.1. Isoprenoids.....	3
1.2.2. Biosynthesis of isoprenoids .....	5
<b>1.3. Non-mevalonate pathway (DXP Pathway) .....</b>	<b>6</b>
<b>1.4. Mevalonate pathway .....</b>	<b>8</b>
<b>1.5. HMG-CoA Reductase .....</b>	<b>10</b>
<b>1.6. Goals for this research .....</b>	<b>12</b>
<b>1.7. Tables and Figures .....</b>	<b>16</b>
<b>1.8. References .....</b>	<b>23</b>
<b>Chapter II: Biophysical characterization of HMGR from <i>Delftia acidovorans</i> and <i>Streptococcus pneumoniae</i>.....</b>	<b>27</b>
<b>2.1. Summary .....</b>	<b>27</b>
<b>2.2. Introduction .....</b>	<b>28</b>
2.2.1. Overall three-dimensional fold of class II HMGR .....	30
2.2.2. Higher order oligomerization.....	32
2.2.3. Goals of this section.....	39
<b>2.3. Results .....</b>	<b>40</b>
2.3.1. Protein Expression and Purification of DaHMGR and SpHMGR.....	40
2.3.2. Overall structure of DaHMGR.....	41
2.3.3. Size Exclusion Chromatography (SEC).....	42
2.3.4. DaHMGR SEC-Small-angle X-ray scattering .....	42
<b>2.4. Discussion .....</b>	<b>44</b>
<b>2.5. Future work .....</b>	<b>50</b>
2.5.1. Stability of DaHMGR.....	50
2.5.2. Oligomerization and activity.....	51
<b>2.6. Materials and methods.....</b>	<b>52</b>
2.6.1. Isolating DaHMGR and SpHMGR.....	52
2.6.2. Crystallization .....	55
2.6.3. Small-angle X-ray Scattering.....	56



2.6.4.	Size Exclusion Chromatography .....	57
<b>2.7.</b>	<b>Tables and figures .....</b>	<b>58</b>
<b>2.8.</b>	<b>References .....</b>	<b>79</b>
<b>Chapter III: Kinetic, and structural characterization of HMGR from <i>Delftia acidovorans</i> and <i>Streptococcus pneumoniae</i> .....</b>		<b>81</b>
<b>3.1.</b>	<b>Summary .....</b>	<b>81</b>
<b>3.2.</b>	<b>Introduction .....</b>	<b>82</b>
3.2.1.	Active site architecture and mechanism .....	84
3.2.2.	Role of the C-terminal domain in catalysis .....	86
3.2.3.	Goals of this section .....	89
<b>3.3.</b>	<b>Results .....</b>	<b>91</b>
3.3.1.	Activity assays .....	91
3.3.2.	Crystal structures .....	92
<b>3.4.</b>	<b>Discussion .....</b>	<b>96</b>
<b>3.5.</b>	<b>Future work .....</b>	<b>107</b>
3.5.1.	Additional kinetic characterization .....	107
3.5.2.	Additional structures with different ligands bound .....	108
3.5.3.	Understanding the forces that enable CTD motions .....	108
<b>3.6.</b>	<b>Materials and methods .....</b>	<b>109</b>
3.6.2.	Citrate inhibition assays .....	110
3.6.3.	Crystallization .....	110
<b>3.7.</b>	<b>Tables and figures .....</b>	<b>115</b>
<b>3.8.</b>	<b>References .....</b>	<b>143</b>
<b>Chapter IV. Altering the cofactor specificity of class II HMGR using rational design.....</b>		<b>146</b>
<b>4.1.</b>	<b>Summary .....</b>	<b>146</b>
<b>4.2.</b>	<b>Introduction .....</b>	<b>147</b>
4.2.1.	NAD(H) and NADP(H) .....	147
4.2.2.	Structural motifs that bind NAD(P)(H) .....	149
4.2.3.	Structural motifs found in class II HMGRs .....	149
4.2.4.	Goals of this section .....	154
<b>4.3.</b>	<b>Results .....</b>	<b>156</b>
4.3.1.	Kinetics .....	156
4.3.2.	X-ray crystal structures .....	159
<b>4.4.</b>	<b>Discussion .....</b>	<b>163</b>
<b>4.5.</b>	<b>Future work .....</b>	<b>171</b>

4.5.1.	Improving overall catalytic activity .....	171
4.5.2.	Investigating the structural features that facilitate the cofactor binding in the upper position in SpHMGR constructs.....	172
4.5.3.	Translating and optimizing the helices between additional class II HMGRs.....	173
<b>4.6.</b>	<b>Materials and methods.....</b>	<b>174</b>
4.6.1.	Designing DaHMGR and SpHMGR mutants.....	174
4.6.2.	Expression and purification of DaHMGR and SpHMGR mutants .....	174
4.6.3.	Kinetic Characterization .....	175
4.6.4.	Crystallization of DaHMGR and SpHMGR mutants.....	176
4.6.5.	X-ray data collection, structure determination, and refinement of DaHMGR and SpHMGR mutants .....	177
<b>4.7.</b>	<b>Tables and figures .....</b>	<b>178</b>
<b>4.8.</b>	<b>References .....</b>	<b>197</b>
<b>Chapter V: Applications for HMGR in natural product, drug design, and metabolic engineering research.....</b>		<b>199</b>
<b>5.1.</b>	<b>Summary .....</b>	<b>199</b>
<b>5.2.</b>	<b>Introduction .....</b>	<b>200</b>
<b>5.3.</b>	<b>HMGR and natural products.....</b>	<b>201</b>
5.3.1.	Why are natural products important?.....	201
5.3.2.	What is the role of HMGR-dependent reactions in natural product research? .....	203
<b>5.4.</b>	<b>HMGR and metabolic engineering .....</b>	<b>207</b>
5.4.1.	What does metabolic engineering offer? .....	207
5.4.2.	How can we exploit intrinsic HMGR activity?.....	209
<b>5.5.</b>	<b>HMGR and drug design .....</b>	<b>217</b>
5.5.1.	What makes a good drug?.....	217
5.5.2.	How can we target class II HMGRs at the active site?.....	219
<b>5.6.</b>	<b>Conclusion.....</b>	<b>228</b>
<b>5.7.</b>	<b>Tables and Figures .....</b>	<b>229</b>
<b>5.8.</b>	<b>References .....</b>	<b>237</b>

## **Chapter I: HMG-CoA reductase, a key enzyme in natural product synthesis**

### **1.1. Natural Products**

An exciting field of research at the interface between chemistry and biology seeks to elucidate and modify the structures of natural products to yield chemicals that have desired functionality and utility across industry and medicine [1]–[10]. These natural products are biological molecules that are produced in nature by living organisms to have some benefit in the context of the respective organism. In their native environments, these compounds typically serve as secondary metabolites that encompass a wide variety of physiological roles including functioning as signaling molecules, such as hormones, or redox molecules in the form of electron acceptors and donors [11]. When extracted, these compounds can also be used as fuels, fragrances, flavors, medicine and in many other industries [12]. With such incredible diversity available on our planet, nature has provided a seemingly endless reservoir of compounds to mine, interrogate, and employ.

These natural product molecules offer unique chemical diversity and bioactivity [13] and it is not surprising that the functions of natural products are distinctly a result of their atomic configurations. To this end, researchers can, for example, exploit specific chemical functional groups that facilitate interactions between a drug and an enzyme, to improve drug efficacy. This structure and function interdependence, therefore, equips researchers with the ability to modify or develop certain features in chemical structures giving rise to molecules with novel applications, a venture that is both scientifically inspiring and practically useful. As an example, in some natural products, such as some heterocyclic alkaloids, the incorporation of nitrogen into its ring moiety

gives the molecule high biological activity [14] as observed in cases such as or quinine, an anti-parasitic medicine used to treat malaria.

While interest in natural products and their derivatives stems from the ability of these compounds to constitute important chemicals such as medicine, nutrition, fragrances, and biofuels, its aspirations are mediated by the success of its research and development. In this study, two major areas of research that greatly benefit from such endeavors are discussed in Chapter Five and involve its applications in drug development and metabolic engineering of commodity chemicals. A prerequisite to being able to utilize these approaches, however, is for researchers to not only know and understand the structure and function of these compounds of interest, but also to explore efficient and cost-effective ways to produce them. The present study, therefore, aspires towards expanding this knowledgebase to further the research and development in the above endeavors.

## **1.2. Isoprenoid Biosynthesis**

The field of natural product biosynthesis, whether it be for the benefit of medicine or to produce other compounds of interest, is an attractive area of study and underscores the importance of research aiming to shed light on the biosynthesis of these important biomolecules. An abundant source of natural products is comprised of the isoprenoids, which compose the largest and most structurally diverse class of natural products [11] and, unsurprisingly, these ubiquitous molecules constitute a wide variety of functions as a result of their vast assortment of structures.

### 1.2.1. Isoprenoids

Even though they are often considered secondary metabolites, in contrast to primary metabolites whose roles are critical to survival, isoprenoids are often essential to the organism [15]. Table 1 provides the structures of a variety of isoprenoids and isoprenoid derivatives along with their functions. What is most remarkable, however, is that despite the extensive diversity of the structures of isoprenoids, all are built with the same isomeric five-carbon building blocks: isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP).

The discovery that all isoprenoids are made from IPP and DMAPP is attributed to a German chemist Otto Wallach who was the first person to characterize isoprenoids, called terpenes in the late 1800's and early 1900's. By studying simple ten-carbon ( $C_{10}$ ) monoterpenes, he began to develop the concept of an isoprene unit as a building block towards larger complex molecules [16]. His advances in terpene chemistry and its applications, won him the Nobel Prize in Chemistry in 1910 [16]. Research on terpenes was further developed by Leopold Ruzicka, a Croatian scientist who also earned the Nobel Prize in Chemistry in 1939, who researched higher-order terpenes and further developed the isoprene rule [16], [17]. This rule outlined the formation of isoprenoids through condensation of  $C_5$  isoprene units in a head-to-tail fashion [17].

Thanks to the above pioneers in this field, we now have a clearer understanding of how polymerization of isoprene units occurs. As shown in the example in Figure 1A, DMAPP and IPP can be condensed in a head-to-tail fashion to give a  $C_{10}$  intermediate, geranyl pyrophosphate (GPP). Further polymerization using additional IPP building blocks can continue in the same manner giving longer chain products, farnesyl pyrophosphate (FPP,  $C_{15}$ ) or geranylgeranyl pyrophosphate (GGPP,  $C_{20}$ ). The steps of polymerization, as outlined in Figure 1B, include

ionization where the pyrophosphate leaves, carbon-carbon bond formation between IPP and DMAPP, and elimination where deprotonation of a neighboring hydrogen neutralizes the carbocation by the formation of a double bond.

These  $C_x$  intermediates can further undergo cyclization by terpene synthases via carbocation chemistry, as shown in Figure 1B, to give drastically different  $C_x$  terpenes, all of which can be additionally modified using downstream enzymes giving rise to tens of thousands of unique isoprenoid structures; this explains the diverse structures and far-reaching applications of these molecules. There is also intrinsic reactivity within the carbocation intermediates that are generated.

While there are many modifying enzymes, this intrinsic reactivity affords isoprenoid molecules the ability to undergo structural rearrangements without the need for additional enzymes beyond the terpene cyclases. Hydride shifts, for example, would enable the rearrangement of the positive charge on the molecule by the migration of a hydrogen and its electrons to the carbocation which results in a new location of the carbocation thus giving the molecule different reactivity. A methyl shift, as the name suggests, involves the migration of a methyl group (as a methyl anion,  $CH_3^-$ ) to a nearby carbon; this step is often coupled with hydride migrations. Pyrophosphate migration is like a hydride migration in that its movement to a new location essentially moves the carbocation with new reactivity.

The carbocation, in most cases, gives the molecule many opportunities for rearrangements and when combined with finely tuned active sites of enzymes can generate a multitude of molecules. The positive charge generated as well as the nature of the terpene cyclase active site that contains the carbocation drive this phenomenon; the distribution of charges and the steric considerations of the active site dictate the outcome of the structural rearrangements, and the

ability to undergo these reactions are intrinsic to the intermediate. As mentioned, the fascinating feature of this extensive class of molecules is that despite their diversity in structure, they all begin with IPP and DMAPP as the foundational building blocks, which brings into focus the biosynthesis of IPP and DMAPP.

### 1.2.2. Biosynthesis of isoprenoids

In the three domains of life, archaea, bacteria and eukaryotes, there are two major biosynthetic pathways that produce the isoprenoid building blocks IPP and DMAPP [18]. The non-mevalonate pathway, also known as the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway, is responsible for synthesizing these isoprenoid building blocks from glycolytic intermediates whereas the mevalonate (MEV) pathway utilizes acetyl-CoA; thus, as shown in Figure 2, the DXP and MEV pathways both utilize substrates from central metabolism. As a point of distinction, these pathways are independent pathways where the DXP pathway is found primarily in bacteria whereas in eukaryotes and archaea the MEV pathway is responsible for producing IPP and DMAPP; however there exists exceptions to this classification [18]. In some instances, for example in plants, both pathways are expressed but differ in their localization, where the MEV pathway is found in the cytosol and the DXP pathway in plastids [18].

Further emphasized in Figure 2, throughout biology the paths towards the biosynthesis of the important isoprenoid precursor, the isoprene unit, is limited to just two metabolic routes, which is discussed below, and as such these pathways continue to inspire research that aims to understand and augment these systems with a proven history of yielding important biomolecules.

### 1.3. Non-mevalonate pathway (DXP Pathway)

While the mevalonate (MEV) pathways it the focus of this study, it is worthwhile to obtain a brief overview of the alternative DXP pathway especially since there are some parallels between the DXP and MEV pathways. Early  $^{13}\text{C}$ -labeling experiments observed in bacteria showed that the synthesis of certain isoprenoids was a result of a different route other than the mevalonate pathway [19]. It would have otherwise been a simple assignment of the atomic arrangement of the molecule using the isoprene unit, but the results suggested an alternative pathway. These  $^{13}\text{C}$ -labeling experiments showed that the isoprene units that were produced were labeled from glycolytic intermediates and subsequently this finding lead researchers to distinguish the starting substrates of the DXP pathway as d-glyceraldehyde 3-phosphate (G3P) and pyruvate and not acetyl coenzyme A (acetyl-CoA) which is the starting substrate in the MEV pathway.

The first step in the DXP pathway begins with the condensation of pyruvate and (G3P) which is catalyzed by 1-deoxy-D-xylulose-5-phosphate (DXP) synthase using thiamine pyrophosphate as a cofactor and yields DXP. In step two, DXP reductoisomerase takes DXP and converts it to 2-C-methylerythritol 4-phosphate (MEP) using nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor and required divalent metal. This step is one that is heavily studied and is a target for fosmidomycin [20] which is an antibiotic that can be used in the treatment of malaria. Subsequently, MEP reacts with the nucleotide cytidine triphosphate (CTP) to produce 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME) with the release of pyrophosphate by the enzyme 4-diphosphocytidyl-2C-methyl-D-erythritol cytidyltransferase.

In step four, an ATP-dependent kinase, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, takes CDP-ME and converts it to 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate (CDP-MEP) by adding a phosphate group to a terminal hydroxyl group. The next



enzyme, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, requires two divalent metals, one of which is always a  $Zn^{2+}$ , to help to align and polarize the substrate phosphate groups to cleave off CMP and yield 2C-methyl-D-erythritol-2,4-cyclodiphosphate (MEcPP) [20]. In the sixth step, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase catalyzes the reduction of MEcPP to 4-hydroxy-3-methyl-2-(E)-butenyl-4-di phosphate (HMBDP) and requires electrons which are believed to come from the [4Fe-4S] clusters of ferredoxin [21]. Another reduction step occurs afterwards using 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, an iron-sulfur protein, and NADPH as a redox cofactor to yield mostly IPP and some DMAPP. Isopentenyl-diphosphate isomerase can interchange IPP and DMAPP to keep the ratio of these molecules suitable for downstream reactions. The last two steps are the least understood with alternative players, such as flavodoxin reductase and flavodoxin, being implicated in some organisms [20]. An overview of the DXP pathway is outlined in Figure 3.

In general, the importance of the DXP pathway is like that of the mevalonate pathway. Since it is the obligate route to isoprenoids in many species, and since isoprenoid molecules have often been shown to be essential for the survival of some organisms, this pathway is an attractive target for drug development of broad spectrum antimicrobial drugs targeting serious diseases, including malaria and TB [20]. Of particular importance is the fact that the DXP pathway is absent in humans, so the ability to discriminate between isoprenoid synthesis in human versus a DXP-dependent pathogen is a promising area of study.

#### 1.4. Mevalonate pathway

The mevalonate (MEV) pathway is the second pathway that is responsible for the biosynthesis of IPP and DMAPP. The pathway's starting substrate, acetyl-CoA, plays a central role in the overall metabolism of the cell, and as a result, the MEV pathway is intimately integrated into broader metabolism. The MEV pathway, therefore, is an important point of regulation especially as it relates to the production of signaling molecules, such as hormones, that support the homeostasis of the organism. For example, in a recent study, the MEV pathway was implicated in the progression of certain cancers. In these cancers, including leukemia, lymphoma, breast and prostate cancers, the MEV pathway was previously shown to be upregulated, and this study proposed the disruption of the MEV pathway in order to disrupt specific essential cellular functions thereby reducing unwanted cell proliferation [22].

An overview of the MEV pathway [23] is presented in Figure 4. Starting with three equivalents of acetyl-CoA, the MEV pathway uses seven enzymes to yield IPP and DMAPP. In step 1, acetoacetyl-CoA thiolase condenses two molecules of acetyl-CoA, yielding acetoacetyl-CoA, after which 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG)-CoA synthase (HMGS) then adds on another acetyl-CoA, resulting in HMG-CoA. These first two steps are not unique to the mevalonate pathway; for example, these steps are utilized in ketogenesis which uses HMG-CoA to eventually create ketone bodies that the body can utilize as a form of energy. Step 3 of the MEV pathway is catalyzed by HMG-CoA reductase (HMGR), which uses two equivalents of NAD(P)H to reduce HMG-CoA to mevalonate with loss of coenzyme A. This HMGR-catalyzed synthesis of mevalonate is the first committed step of the MEV pathway, which is how the pathway gets its name. The HMGR-catalyzed reaction is also the rate determining step of the MEV pathway. In Steps 4 and 5, mevalonate-5-kinase (MK) and 5-phosphomevalonate kinase (PMK) add two

phosphate groups sequentially from ATP, leaving a pyrophosphate group on the C5 position to first yield mevalonate-5-phosphate and then mevalonate-5-pyrophosphate. In step 6, using one molecule of ATP, mevalonate-5-phosphate decarboxylase (PMD) removes the carboxylate group, giving isopentenyl pyrophosphate (IPP), the first major product of the MEV pathway. An additional step catalyzed by isopentenyl pyrophosphate isomerase (IPPI) interchanges IPP to dimethylallyl pyrophosphate (DMAPP), the second major product of this pathway.

As pointed out in Figure 1, the products of the MEV pathway, IPP and DMAP can polymerize in a head-to-tail fashion, creating longer chain isoprenoid intermediates. Additional enzymes, such as cyclases, can make modifications to further diversify the chemical structure of products made from this pathway. Since it is an obligate pathway towards the building blocks of these diverse structures, the MEV pathway is often studied in the context of biosynthetic platforms. One study utilized clustered regularly interspaced short palindromic repeats (CRISPR) interference to transcriptionally attenuate the translation of enzymes of the MEV pathway for the purpose of increasing isoprenoid production in *Escherichia coli* [24]. Using this modified pathway, the researchers were able to improve the production of (-)- $\alpha$ -bisabolol, a C<sub>15</sub> isoprenoid that is considered to have anti-inflammatory and skin-soothing properties, while optimizing protein quantity and reducing toxic intermediates [24], [25]. For this reason, a greater understanding of each enzyme of the MEV pathway is warranted and the next section will focus on one critical enzyme in this system, HMG-CoA reductase.

## 1.5. HMG-CoA Reductase

HMGR catalyzes the rate-limiting step of the MEV pathway, converting HMG-CoA to mevalonate using two equivalents of reducing cofactors nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADPH) as shown in Figure 4 [26]. Figure 5 shows a detailed view of this four-electron redox reaction, outlining the substrate and cofactors involved, the intermediates, and the products of the reaction that HMGR catalyzes.

Genes encoding HMGR have been found in all three domains of life with more than 150 HMGR homologs recorded that have evolved into two classes: class I HMGR is found in eukaryotes and in some archaea, and class II HMGR is found in some bacteria and archaea [26]. The catalytic domain of class I is comparable to that of class II HMGR; however, class I HMGRs have two additional domains that are not present in class II enzymes, namely the transmembrane domain and the linker, which influence the localization of HMGR such that class I enzymes are membrane-bound, with the catalytic domain being cytosolic [27]. Since class II HMGRs lack the transmembrane and linker domains, they exist in the cytosol [27]. When comparing the catalytic domain of class I with class II HMGRs, the overall structures look similar. Figure 6 shows the comparison between the overall folds of class I HMGR and class II HMGR, using human and *Pseudomonas mevalonii* HMGR homologs, respectively, and includes a schematic of the domains found in each class of HMGR.

Despite the overall similarity between the structures of the catalytic domain of class I and the class II HMGRs, there are several noteworthy differences between the classes. For example, class I HMGRs have an additional loop containing residues that are implicated in catalysis in addition to residues involved in HMG-CoA binding and oligomerization that reveal differences between the classes [28]. Another difference between HMGR classes is that class I HMGRs

exclusively utilize NADPH while class II HMGR cofactor preference varies [29]. In class II enzymes, HMGRs from *Pseudomonas mevalonii* (PmHMGR), *Delftia acidovorans* (DaHMGR), and *Burkholderia cenocepacia* (BcHMGR) all prefer NADH whereas HMGRs from *Enterococcus faecalis* (EfHMGR) and *Streptococcus pneumoniae* (SpHMGR) prefer NADPH [29]–[33]. Other HMGRs, such as those from *Archaeoglobus fulgidus* (AfHMGR) and *Listeria monocytogenes* (LmHMGR) appear to have diminished cofactor selectivity owed to their ability to utilize either cofactor comparably [34], [35]. In spite of these differences, it is thought that their relatively high degree of structural similarity is a result of divergent evolution [28].

Research into both class I and class II enzymes has aimed at targeting HMGR to control downstream product formation. In an example involving class I HMGR, researchers targeted human HMGR with the goal of regulating the biosynthesis of hormones since the MEV pathway generates products that go on to become product hormones. The most well-known application of this is the use of statin drugs to inhibit the function of class I HMGR, thus inhibiting the production of the sterol cholesterol, and therefore establishing itself as a therapeutic target for regulating cholesterol synthesis. In the 1980's, Michael S. Brown and Joseph L. Goldstein, doctors in the Department of Medicine at the University of Texas Southwestern Medical Center, related HMGR inhibition with the downregulation of low-density lipoprotein to treat cholesterol-related cardiovascular disease. This work eventually won the Nobel Prize in 1985, and two years later, the pharmaceutical giant Merck put out the first FDA-approved statin inhibitor, Lovastatin. Since then, statins, whose primary target and mode of action are through its inhibition of HMGR, have become the bestselling class of pharmaceutical compounds of all time, generating tens of billions of dollars in sales each year [36], [37].

In an example that involves class II HMGR, researchers introduced the MEV pathway of *Saccharomyces cerevisiae* into engineered *E. coli* alongside additional enzymes that produce amorphadiene, the isoprenoid precursor to the antimalarial drug artemisinin [31]. In this study, they interchanged the exogenous class I HMGR from *Saccharomyces cerevisiae* (ScHMGR) with several different class II HMGR homologs and observed the resulting titers of amorphadiene. The authors point out that expression of exogenous pathways in this manner can lead to toxic metabolites, cofactor imbalances or protein expression issues, all resulting in lower yields [31]. However, they found a 120% improvement in amorphadiene production, which was achieved by varying one factor, the HMGR homolog, is a notable achievement considering the breadth of possible complications one may confront in engineering heterologous expression systems.

## **1.6. Goals for this research**

So far, this chapter has introduced the importance of natural products in medicine and industry, and homed in on isoprenoids, the largest and most diverse class of natural products. This led to an overview of the two obligate pathways that produce the isoprene building blocks of isoprenoids, with an emphasis on the mevalonate pathway. Within the mevalonate pathway, HMGR, which catalyzes the first-committed step, was highlighted and in doing so put into context this biologically important enzyme. Specifically, this introduction serves to underscore its value as a central enzyme in the mevalonate pathway, a pathway that is often used and engineered to generate the precursors for some of nature's most important molecules.

Our focus now pivots towards the in-depth study of class II HMGRs which is the core of the exploration presented in this research. In a broad sense, the vision for this study is to advance

our knowledge into the mechanism of catalysis of class II HMGR in order to further characterize this key enzyme. Using a variety of biochemical techniques that survey structural and functional characteristics, and by investigating several different class II HMGR homologs, this work aims to elucidate and interrogate HMGR catalysis especially in the areas of cofactor specificity and structural mechanics.

In Chapter Two, a variety of biophysical experiments are described using class II HMGRs from *Delftia acidovorans* (DaHMGR) and *Streptococcus pneumoniae* (SpHMGR). This will be prefaced with a survey of the literature on class II HMGR with regards to how HMGR presents itself in its native state among class II HMGRs. Topics such as oligomerization and the overall three-dimensional fold will be investigated to fully appreciate the experimental findings revealed in this chapter. Results will be described from techniques involved in isolating DaHMGR and SpHMGR, including expressing and purification methods, as well as biophysical characteristics obtained from small angle X-ray scattering (SAXS), analytical ultracentrifugation (AUC), and size exclusion chromatography (SEC). X-ray crystallography will be also used to substantiate the overall structural features.

In Chapter Three, we will explore the kinetic behavior of a previously uncharacterized HMGR, DaHMGR, and we will complement these studies with atomic-level insight into DaHMGR catalysis. We will consult the literature on class II HMGR with regards to the enzyme mechanism, HMG-CoA and cofactor affinity, additional Michaelis-Menten kinetic parameters, and an overview of the crystal structures that have been published thus far. The results presented in this chapter include those from enzyme activity assays utilizing steady-state kinetics experiments and X-ray crystallographic structures comparing apo and ligand bound structures. Special attention will be given to the C-terminal domain (CTD) that facilitates enzyme catalysis

through the analysis of novel structures of DaHMGR and additional crystal structures obtained from SpHMGR mutagenesis studies showing the CTD in various conformations.

In Chapter Four, we will explore the structural features that govern cofactor specificity among several class II HMGRs. To set this up, there will be a survey of the literature on class II HMGR with regards to cofactor specificity and the structural motifs that govern these interactions. Much attention will be dedicated towards an  $\alpha$ -helix that has been shown to dictate cofactor preference, which we term the “cofactor helix” of HMGR. Utilizing the knowledge gained from the prior chapters, protein engineering will be utilized to generate modified HMGRs with altered cofactor specificity. The results presented will involve enzyme activity assays and crystal structures with cofactors bound of these modified HMGRs and will demonstrate the power of the complementary nature of the function of these enzymes and the structure, on a residue-by-residue basis, as well as on a secondary structure basis.

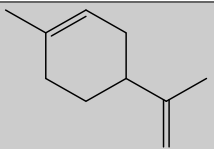
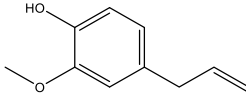
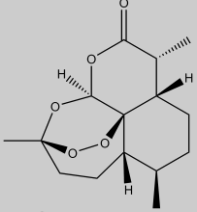
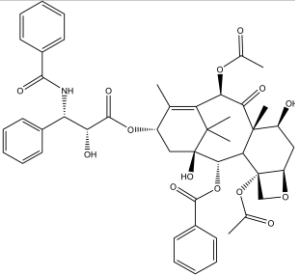
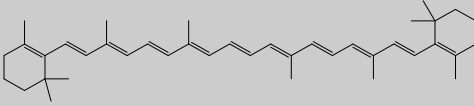
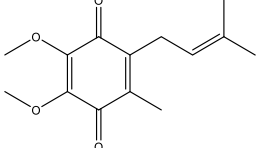
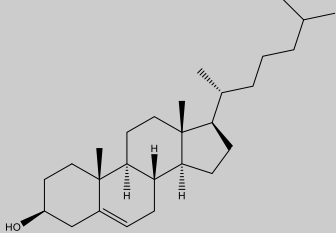
In the final chapter, Chapter Five, the potential applications of this research will be explored, especially in metabolic engineering in HMGR-dependent natural product synthesis, and drug development targeting HMGR. One of the hurdles of this exciting field is the need to overcome bottlenecks to improve product formation [12]. When bioengineering microbial hosts to produce desired products, whether it is for medicine or for other chemicals of interest, cofactor requirements can be a source of bottlenecks [4]. To this end, knowing an enzyme’s specificity for a cofactor is crucial in providing the optimal functioning of the pathway to prevent delay in the intermediate steps. It is also conceivable that with sufficient knowledge of the basis of cofactor affinity and specificity, one could alter this cofactor preference to either make the enzyme more promiscuous for both NADPH and NADH, or to improve the affinity of its preferred cofactor and therefore lower its bulk concentration requirement. In addition, with atomic-level insight of the

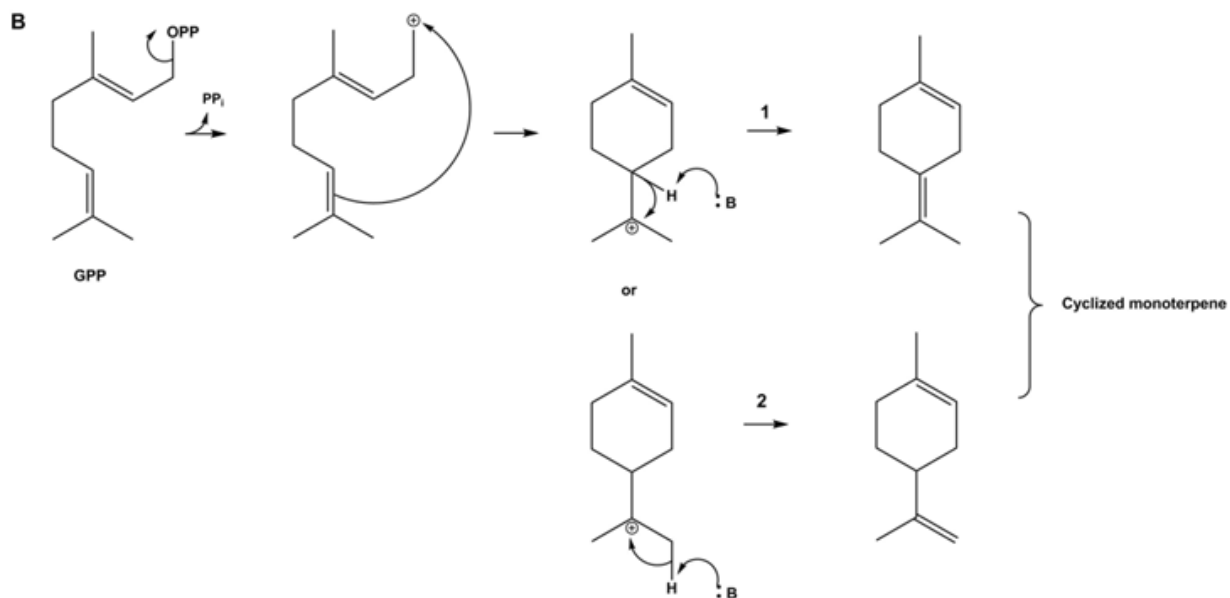
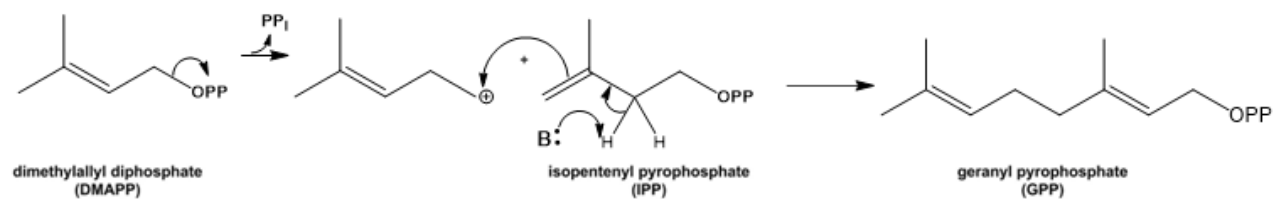


active site architecture of specific HMGRs, antibiotics with high specificity can be developed to target HMGRs in specific pathogens. In sum, these endeavors make the study presented here an important resource in the promising field of HMGR-dependent isoprenoid natural product biosynthesis and class II HMGR antibiotic development.

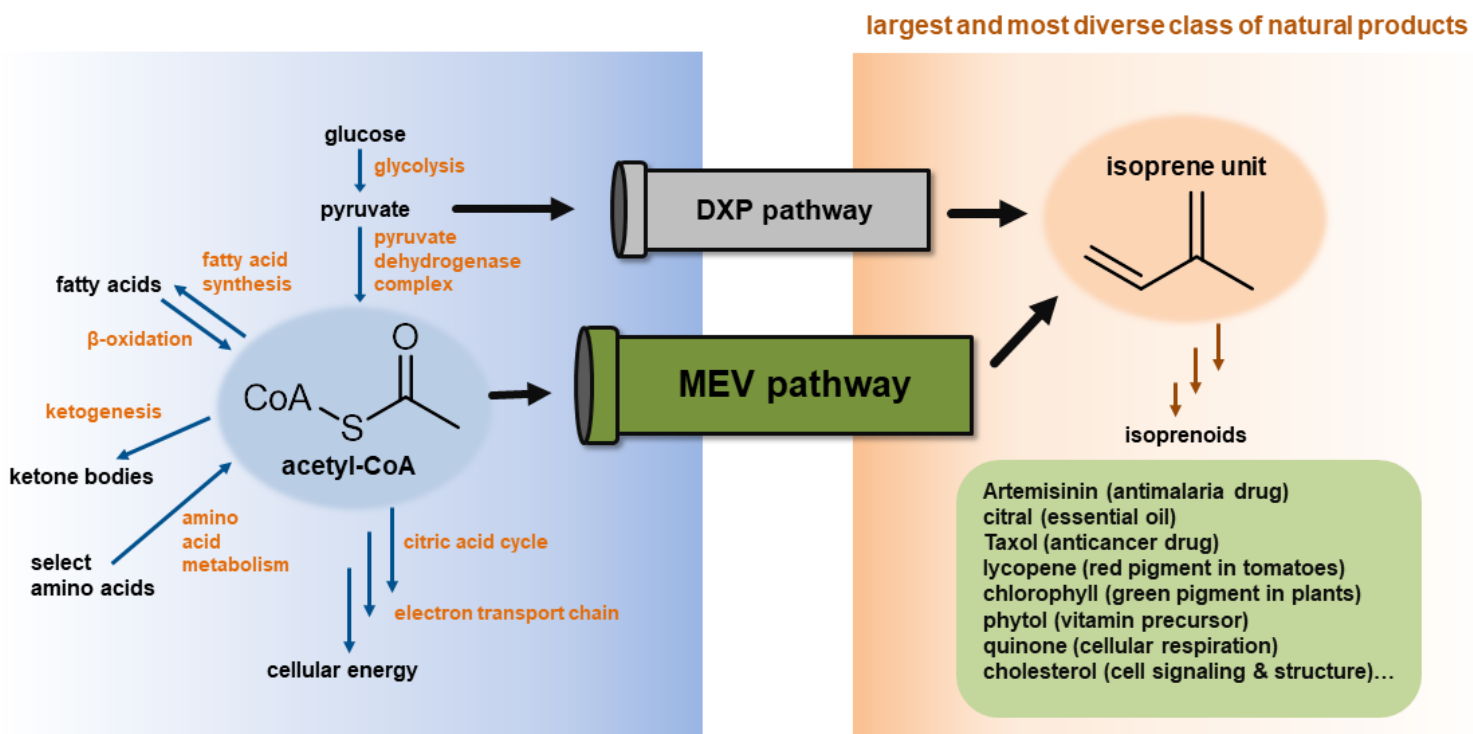
## 1.7. Tables and Figures

**Table 1. Examples of isoprenoids.** A few isoprenoid derivatives, their structures, and their functions, illustrating the diversity within these molecules.

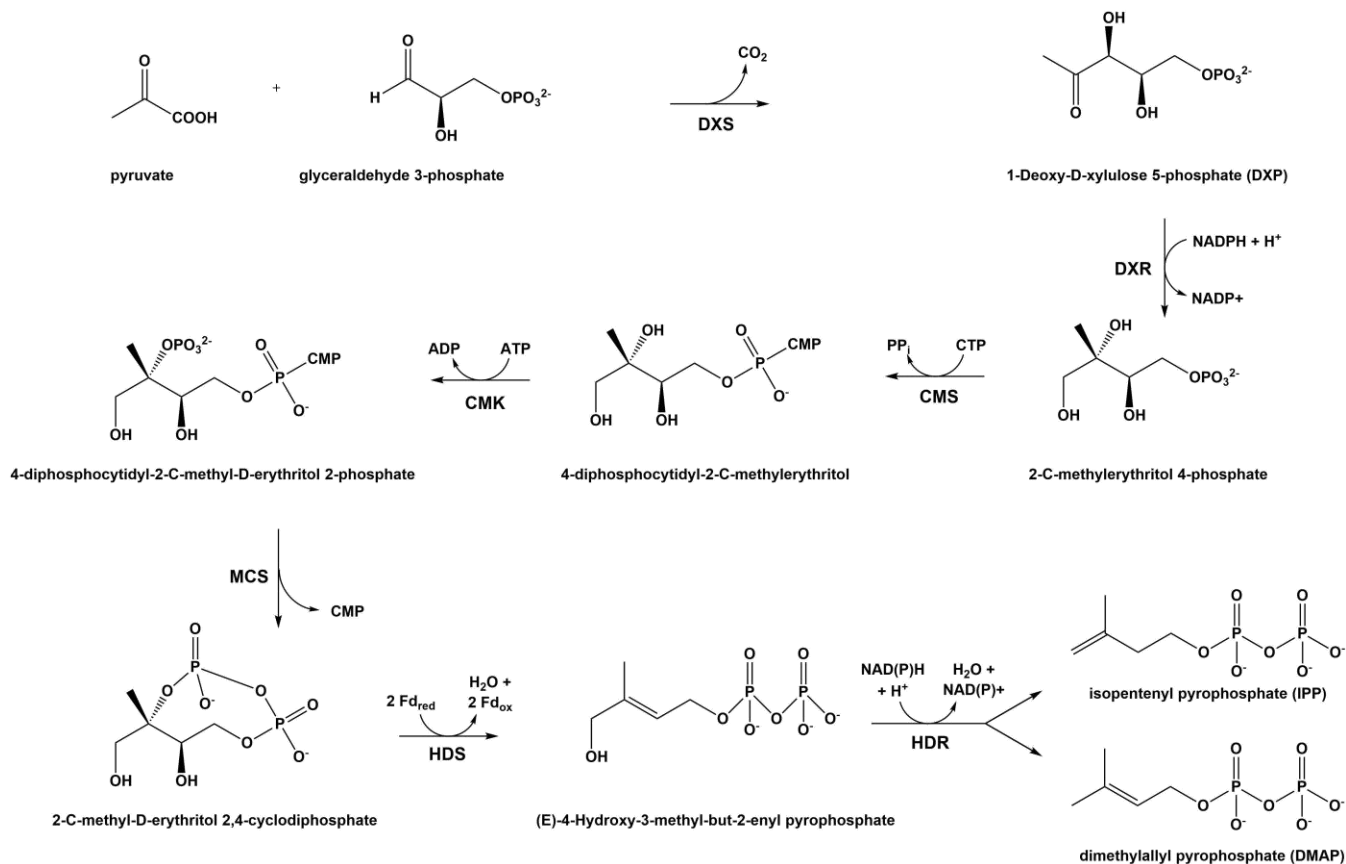
Compound	Structure	Function
<b>Limonene,</b> [38]		<ul style="list-style-type: none"> <li>- oil with chemo-preventative and antitumor activity</li> <li>- fragrance, flavor, and cosmetics</li> <li>-</li> </ul>
<b>eugenol,</b> [39]		<ul style="list-style-type: none"> <li>- clove oil, used in food</li> <li>- herbal benefits for use related to gastrointestinal, respiratory</li> <li>- topical alleviation of toothache</li> </ul>
<b>artemisinin,</b> [40]		<ul style="list-style-type: none"> <li>- antimalarial</li> <li>- treatment of multi-drug resistant strains of malaria</li> </ul>
<b>paclitaxel</b> (Taxol), [41]		<ul style="list-style-type: none"> <li>- cancer drug</li> <li>- treats breast, lung and ovarian cancer</li> <li>- used in treatment of Kaposi's sarcoma</li> </ul>
<b>beta-Carotene,</b> [42]		<ul style="list-style-type: none"> <li>- vitamin A precursor</li> <li>- antineoplastic and chemo-preventive activity</li> <li>- antioxidant</li> <li>- induces cell differentiation and apoptosis of some tumor cell types</li> </ul>
<b>ubiquinone-1,</b> [43]		<ul style="list-style-type: none"> <li>- involved in mitochondrial electron transport chain</li> <li>- potent antioxidant</li> </ul>
<b>cholesterol,</b> [44]		<ul style="list-style-type: none"> <li>- part of cell membranes</li> <li>- vitamin D and steroid precursor</li> <li>- roles in brain synapses and immune system</li> </ul>



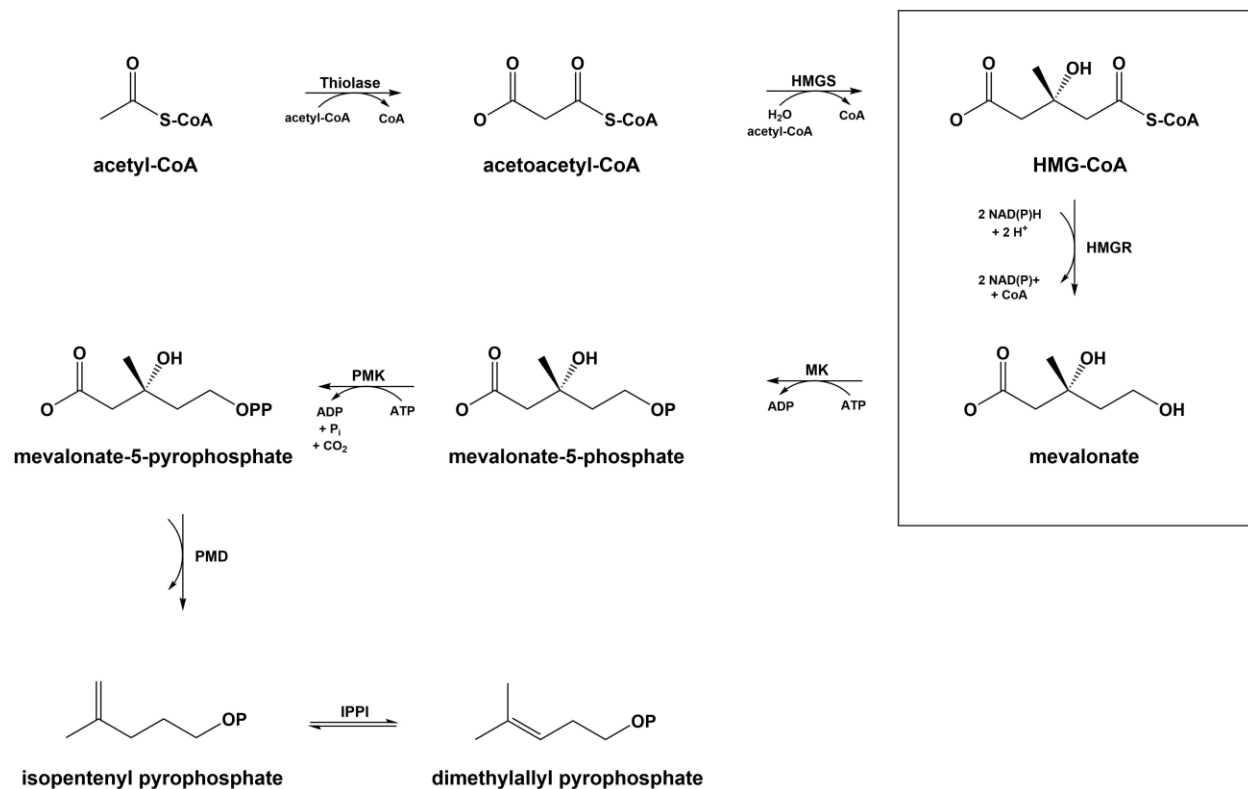
**Figure 1. Condensation of IPP and DMAPP building blocks.** (A) DMAPP and IPP join in head-to-tail fashion, yielding a monoterpene, GPP. (B) GPP is rearranged into cyclized monoterpenes. The difference between reaction 1 and 2 in Figure 3B, is dictated by the site of deprotonation by a basic sidechain.



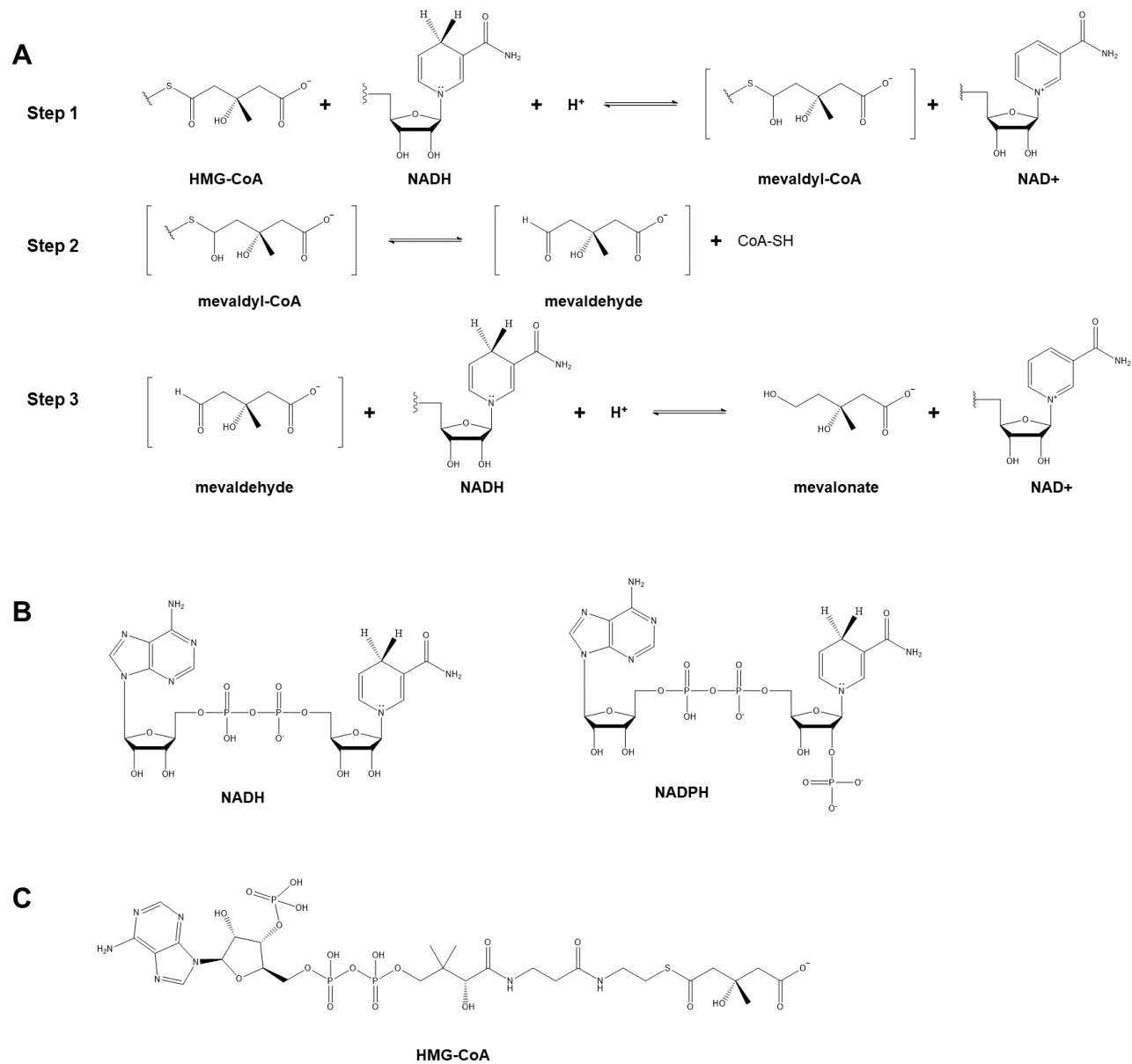
**Figure 2. From acetyl-CoA to isoprenoids.** Acetyl-CoA is implicated in most of the important metabolic pathways of the cell as illustrated above. There are, however, two major pathways that convert lead to isoprene units, which are the building blocks for isoprenoids, the largest and most diverse class of natural products. The mevalonate and DXP pathways are, therefore, crucial metabolic pathways in the synthesis of a wide range of important biomolecules.



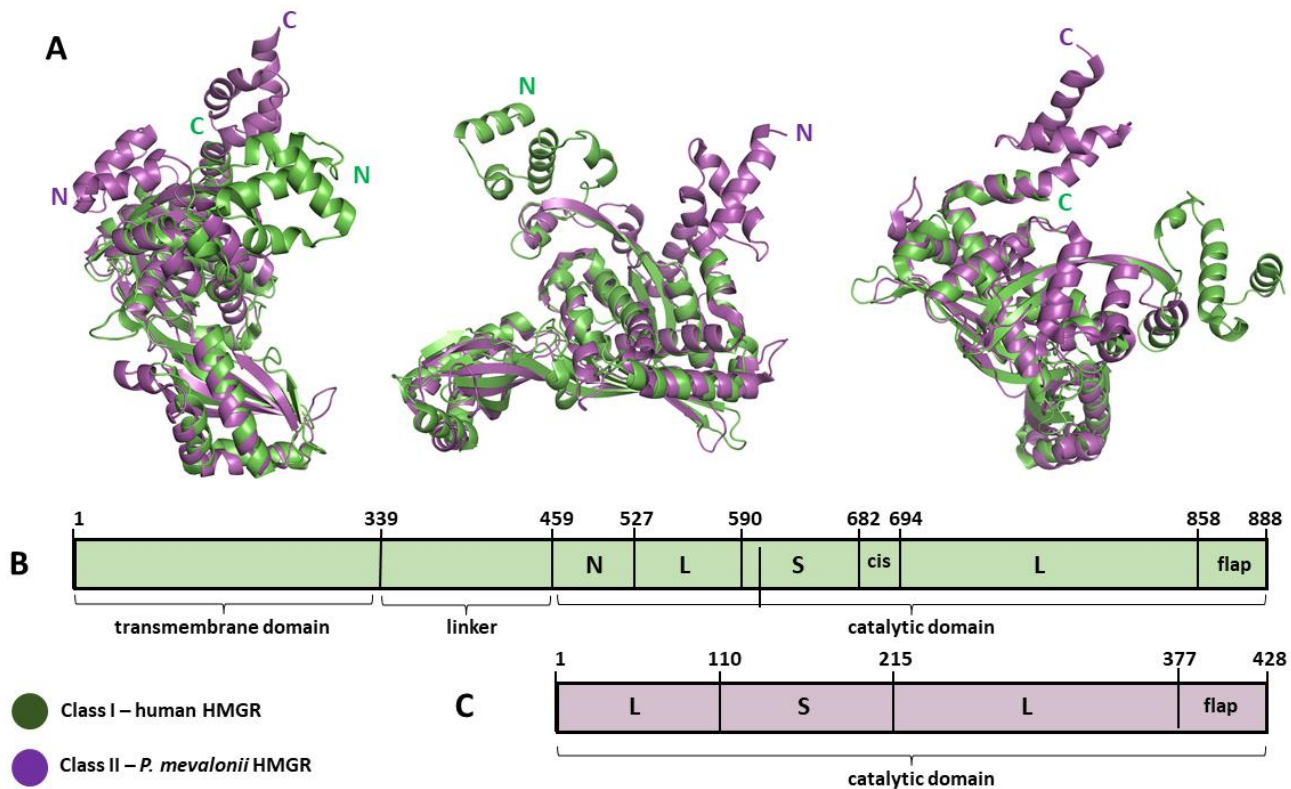
**Figure 3. DXP pathway.** An overview of the steps of the non-mevalonate pathway, often referred to as the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway.



**Figure 4. Overview of the mevalonate pathway.** Six enzymes convert acetyl-CoA to isopentenyl pyrophosphate (IPP), with an additional enzyme interconverting IPP to DMAPP. P represents a phosphate group. HMG-R, the focus of this research, performs the reaction that is outlined in the box on the right, converting HMG-CoA to mevalonate using 2 NAD(P)H.



**Figure 5. Overview of the HMGR reaction.** A) Basic step-by-step overview of the reaction that HMG-CoA reductase catalyzes. The precise order of steps remains unknown. NADH is used as a representative cofactor, although cofactor usage among HMGRs vary. The entire molecules of B) NADH and NADPH, and C) HMG-CoA are shown whereas it is truncated in A.



**Figure 6. Comparison between the structures class I and class II HMGR.** A) Various views of an overlay of the catalytic portions of a class I HMGR (human HMGR in green PDB 1DQA) and a class II HMGR (*P. mevalonii* HMGR in purple PDB 1QAX) with the N and C termini indicated. B) and C) show primary structure schematics with residue numbering indicated of the various domains of human HMGR and *P. mevalonii* HMGR, respectively, adapted from Friesen J., Rodwell V., *Genome Biology*. (2004). The domains are denoted as follows: N = N domain, found in human HMGR, that connects the linker to the large domain; L = large domain; S = small domain; cis = region that contains the cis-loop that facilitates HMG-CoA and NADP(H) binding, flap = mobile C-terminal domain. As depicted in the schematic, the catalytic domains are comparable, and class II HMGRs lack the linker and transmembrane domains. The large domain includes the region that binds HMG-CoA, and the small domain contains the region that binds NAD(P)H.



## 1.8. References

- [1] G. Karageorgis, D. J. Foley, L. Laraia, and H. Waldmann, “Principle and design of pseudo-natural products,” *Nat. Chem.*, vol. 12, no. 3, pp. 227–235, Mar. 2020, doi: 10.1038/s41557-019-0411-x.
- [2] J. W.-H. Li and J. C. Vederas, “Drug Discovery and Natural Products: End of an Era or an Endless Frontier?,” *Science*, vol. 325, no. 5937, pp. 161–165, Jul. 2009, doi: 10.1126/science.1168243.
- [3] H. van Hattum and H. Waldmann, “Biology-Oriented Synthesis: Harnessing the Power of Evolution,” *J. Am. Chem. Soc.*, vol. 136, no. 34, pp. 11853–11859, Aug. 2014, doi: 10.1021/ja505861d.
- [4] M. M. M. Pinto *et al.*, “From Natural Products to New Synthetic Small Molecules: A Journey through the World of Xanthonones,” *Molecules*, vol. 26, no. 2, p. 431, Jan. 2021, doi: 10.3390/molecules26020431.
- [5] C. Lyu *et al.*, “CMNPD: a comprehensive marine natural products database towards facilitating drug discovery from the ocean,” *Nucleic Acids Research*, vol. 49, no. D1, pp. D509–D515, Jan. 2021, doi: 10.1093/nar/gkaa763.
- [6] L. Barra and I. Abe, “Chemistry of fungal meroterpenoid cyclases,” *Nat. Prod. Rep.*, vol. 38, no. 3, pp. 566–585, 2021, doi: 10.1039/D0NP00056F.
- [7] Z.-J. Ni *et al.*, “Recent updates on the chemistry, bioactivities, mode of action, and industrial applications of plant essential oils,” *Trends in Food Science & Technology*, vol. 110, pp. 78–89, Apr. 2021, doi: 10.1016/j.tifs.2021.01.070.
- [8] J. Bachořík and M. Urban, “Biocatalysis in the Chemistry of Lupane Triterpenoids,” *Molecules*, vol. 26, no. 8, p. 2271, Apr. 2021, doi: 10.3390/molecules26082271.
- [9] A. Adhikari, B. Shen, and C. Rader, “Challenges and opportunities to develop enediynes natural products as payloads for antibody-drug conjugates,” *Antibody Therapeutics*, vol. 4, no. 1, pp. 1–15, Jan. 2021, doi: 10.1093/abt/tbab001.
- [10] F. Ghirga *et al.*, “A unique high-diversity natural product collection as a reservoir of new therapeutic leads,” *Org. Chem. Front.*, vol. 8, no. 5, pp. 996–1025, 2021, doi: 10.1039/D0QO01210F.
- [11] B. M. Lange, T. Rujan, W. Martin, and R. Croteau, “Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes,” *Proc Natl Acad Sci USA*, vol. 97, no. 24, p. 13172, Nov. 2000, doi: 10.1073/pnas.240454797.
- [12] J. D. Keasling, “Manufacturing Molecules Through Metabolic Engineering,” *Science*, vol. 330, no. 6009, p. 1355, Dec. 2010, doi: 10.1126/science.1193990.

- [13] S. D. Sarker, Ed., *Natural products isolation*, 3. ed. New York, NY Heidelberg: Humana Press, 2012.
- [14] P. Dey *et al.*, “Analysis of alkaloids (indole alkaloids, isoquinoline alkaloids, tropane alkaloids),” in *Recent Advances in Natural Products Analysis*, Elsevier, 2020, pp. 505–567. doi: 10.1016/B978-0-12-816455-6.00015-9.
- [15] E. I. Wilding *et al.*, “Essentiality, Expression, and Characterization of the Class II 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase of *Staphylococcus aureus*,” *J Bacteriol*, vol. 182, no. 18, pp. 5147–5152, Sep. 2000, doi: 10.1128/JB.182.18.5147-5152.2000.
- [16] M. Christmann, “Otto Wallach: Founder of Terpene Chemistry and Nobel Laureate 1910,” *Angew. Chem. Int. Ed.*, vol. 49, no. 50, pp. 9580–9586, Dec. 2010, doi: 10.1002/anie.201003155.
- [17] S. G. Hillier and R. Lathe, “Terpenes, hormones and life: isoprene rule revisited,” *Journal of Endocrinology*, vol. 242, no. 2, pp. R9–R22, Aug. 2019, doi: 10.1530/JOE-19-0084.
- [18] J. Kirby and J. D. Keasling, “Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering,” *Annu. Rev. Plant Biol.*, vol. 60, no. 1, pp. 335–355, Jun. 2009, doi: 10.1146/annurev.arplant.043008.091955.
- [19] M. Rohmer and M. Rohmer, “The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants†,” *Nat. Prod. Rep.*, vol. 16, no. 5, pp. 565–574, 1999, doi: 10.1039/a709175c.
- [20] W. N. Hunter, “The Non-mevalonate Pathway of Isoprenoid Precursor Biosynthesis,” *Journal of Biological Chemistry*, vol. 282, no. 30, pp. 21573–21577, Jul. 2007, doi: 10.1074/jbc.R700005200.
- [21] K. Okada and T. Hase, “Cyanobacterial Non-mevalonate Pathway,” *Journal of Biological Chemistry*, vol. 280, no. 21, pp. 20672–20679, May 2005, doi: 10.1074/jbc.M500865200.
- [22] F. Tamanoi, M. Azizian, M. Ashrafi, and S. Bathaie, “Mevalonate Pathway and Human Cancers,” *CMP*, vol. 10, no. 2, pp. 77–85, Mar. 2017, doi: 10.2174/1874467209666160112123205.
- [23] W. Eisenreich, A. Bacher, D. Arigoni, and F. Rohdich, “Biosynthesis of isoprenoids via the non-mevalonate pathway,” *CMLS, Cell. Mol. Life Sci.*, vol. 61, no. 12, Jun. 2004, doi: 10.1007/s00018-004-3381-z.
- [24] S. K. Kim *et al.*, “CRISPR interference-guided balancing of a biosynthetic mevalonate pathway increases terpenoid production,” *Metabolic Engineering*, vol. 38, pp. 228–240, Nov. 2016, doi: 10.1016/j.ymben.2016.08.006.
- [25] K. Russell and S. E. Jacob, “Bisabolol,” *Dermatitis*, vol. 21, no. 1, pp. 57–58, Feb. 2010.

- [26] J. A. Friesen and V. W. Rodwell, “The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases,” *Genome Biol*, vol. 5, no. 11, p. 248, 2004, doi: 10.1186/gb-2004-5-11-248.
- [27] D. S. Gesto, C. M. S. Pereira, N. M. F. S. Cerqueira, and S. F. Sousa, “An Atomic-Level Perspective of HMG-CoA-Reductase: The Target Enzyme to Treat Hypercholesterolemia,” *Molecules*, vol. 25, no. 17, p. 3891, Aug. 2020, doi: 10.3390/molecules25173891.
- [28] E. Istvan, “Bacterial and mammalian HMG-CoA reductases: related enzymes with distinct architectures,” *Current Opinion in Structural Biology*, vol. 11, no. 6, pp. 746–751, Dec. 2001, doi: 10.1016/S0959-440X(01)00276-7.
- [29] B. R. Miller and Y. Kung, “Structural Features and Domain Movements Controlling Substrate Binding and Cofactor Specificity in Class II HMG-CoA Reductase,” *Biochemistry*, vol. 57, no. 5, pp. 654–662, Feb. 2018, doi: 10.1021/acs.biochem.7b00999.
- [30] J. A. Friesen, C. M. Lawrence, C. V. Stauffacher, and V. W. Rodwell, “Structural Determinants of Nucleotide Coenzyme Specificity in the Distinctive Dinucleotide Binding Fold of HMG-CoA Reductase from *Pseudomonas mevalonii* †,” *Biochemistry*, vol. 35, no. 37, pp. 11945–11950, Jan. 1996, doi: 10.1021/bi9609937.
- [31] S. M. Ma *et al.*, “Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases,” *Metabolic Engineering*, vol. 13, no. 5, pp. 588–597, Sep. 2011, doi: 10.1016/j.ymben.2011.07.001.
- [32] B. H. Schwarz *et al.*, “Kinetic characterization of an oxidative, cooperative HMG-CoA reductase from *Burkholderia cenocepacia*,” *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1844, no. 2, pp. 457–464, Feb. 2014, doi: 10.1016/j.bbapap.2013.11.015.
- [33] M. Hedl *et al.*, “*Enterococcus faecalis* Acetoacetyl-Coenzyme A Thiolase/3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase, a Dual-Function Protein of Isopentenyl Diphosphate Biosynthesis,” *J Bacteriol*, vol. 184, no. 8, pp. 2116–2122, Apr. 2002, doi: 10.1128/JB.184.8.2116-2122.2002.
- [34] D.-Y. Kim, C. V. Stauffacher, and V. W. Rodwell, “Dual coenzyme specificity of *Archaeoglobus fulgidus* HMG-CoA reductase,” *Protein Sci.*, vol. 9, no. 6, pp. 1226–1234, 2000, doi: 10.1110/ps.9.6.1226.
- [35] A. E. Theivagt, E. N. Amanti, N. J. Beresford, L. Taberner, and J. A. Friesen, “Characterization of an HMG-CoA Reductase from *Listeria monocytogenes* That Exhibits Dual Coenzyme Specificity,” *Biochemistry*, vol. 45, no. 48, pp. 14397–14406, Dec. 2006, doi: 10.1021/bi0614636.
- [36] A. Endo, “A historical perspective on the discovery of statins,” *Proc. Jpn. Acad., Ser. B*, vol. 86, no. 5, pp. 484–493, 2010, doi: 10.2183/pjab.86.484.

- [37] John LaMattina, "The Story Of Statins," *Chemical & Engineering News*, Jun. 08, 2009. Accessed: Jun. 30, 2021. [Online]. Available: <https://cen.acs.org/articles/87/i23/Story-Statins.html>
- [38] National Center for Biotechnology Information, "Limonene," *PubChem*, 2021. <https://pubchem.ncbi.nlm.nih.gov/compound/Limonene> (accessed Sep. 07, 2021).
- [39] National Center for Biotechnology Information, "Eugenol," *PubChem*, 2021. <https://pubchem.ncbi.nlm.nih.gov/compound/3314> (accessed Sep. 07, 2021).
- [40] National Center for Biotechnology Information, "Artemisinin," *PubChem*, 2022. <https://pubchem.ncbi.nlm.nih.gov/compound/Artemisinin> (accessed Mar. 23, 2022).
- [41] "Taxol® (NSC 125973)," *National Cancer Institute*. [https://ntp.cancer.gov/timeline/flash/success\\_stories/S2\\_taxol.htm](https://ntp.cancer.gov/timeline/flash/success_stories/S2_taxol.htm)
- [42] National Center for Biotechnology Information, "beta-Carotene," *PubChem*, 2021. <https://pubchem.ncbi.nlm.nih.gov/compound/5280489> (accessed Sep. 07, 2021).
- [43] National Center for Biotechnology Information, "Ubiquinone-1," *PubChem*, 2021. <https://pubchem.ncbi.nlm.nih.gov/compound/Ubiquinone-1> (accessed Sep. 07, 2021).

## **Chapter II: Biophysical characterization of HMGR from *Delftia acidovorans* and *Streptococcus pneumoniae***

### **2.1. Summary**

Investigating the structural and functional dynamics of an enzyme *in vitro* is often aided by visualization of the protein using techniques such as X-ray crystallography or Cryo-Electron Microscopy and is complemented by functional studies assessing the thermodynamic and kinetic behavior of the protein of interest. While X-ray crystallography can provide atomic-level detail of the structure, which is valuable when making claims about steps of the mechanism, a limitation of this technique is that it requires the creation and optimization of protein crystals that necessarily diffract to high enough resolution. Additionally, X-ray crystallography provides a static snapshot of the protein embedded in the crystal lattice, potentially restricting the dynamic nature of the protein. For this reason, complementary biophysical analysis using alternative techniques must be consulted to substantiate and corroborate the claims being made.

Depending on the nature of the hypothesis, a more comprehensive understanding of the protein must utilize approaches that appropriately addresses how an enzyme behaves in its native environment. Often, this is accomplished *in vitro* by studying the protein in solution. In this chapter, we will review the literature on class II HMGRs as it pertains to the overall structure by examining both the static snapshots provided by X-ray crystallographic analysis as well as dynamic, solution state studies. We will explore some of the methodology of these techniques in detail to recognize the advantages and limitations as we begin to characterize class II HMGR *in vitro*. HMGR proves to be an intriguing enzyme from a structural standpoint due the intricate and

dynamic mechanism which it performs. These experiments were performed using HMGR from the lesser characterized *Delftia acidovorans* (DaHMGR), as well as HMGR from *Streptococcus pneumoniae* (SpHMGR).

Following an overview of the expression and purification of the two HMGRs, X-ray crystallography will outline the overall structural features of these enzymes including the global fold, and secondary structural motifs. Size Exclusion Chromatography (SEC) will reveal that these proteins exist in different oligomerization states. SEC will show that SpHMGR exists as a dimer in solution which is corroborated by the way the protein crystallizes in the crystal lattice. In contrast, DaHMGR in the crystal structure forms a hexamer. However, as revealed by SEC, DaHMGR will be shown to exist in an equilibrium between multiple oligomeric states, requiring small-Angle X-ray Scattering, equipped with an SEC, to assess the oligomeric state of the predominant species, and will suggest that DaHMGR exists in an equilibrium between dimer, and a hexamer. Collectively, these results will begin to lay the foundation as to how class II HMGR functions in the cell and more specifically will begin to address its varying tertiary states within the context of catalysis.

## **2.2. Introduction**

Structure and function are inextricably linked across biology and this relationship is particularly exhibited in protein folding. Thus, decades of research have sought to examine the structure of proteins to visualize how they perform specific, life-giving functions. Meanwhile, parallel work has been done to investigate how proteins, from their primary amino acid sequence, fold into elaborate three-dimensional structures [1]. The most recent advances to this end has

produced a computational tool, called *AlphaFold 2*, that is able to predict structures purely from their amino acid sequence by utilizing machine learning that surveys physical and biological data as part of its algorithm – a remarkable achievement [2]!

Among other factors, these computational methods consider the various protein folding forces that facilitate the formation of its three-dimensional conformation. Noncovalent interactions such as hydrophobicity, hydrogen bonding, Coulombic interactions, and van der Waals interactions are at play and collectively aid in the thermodynamics of achieving a more stable conformation [3]. There are also secondary structural interactions between the backbone of the polypeptide chain, and further interactions between the side chains involved [3]. These interactions not only facilitate the formation of the tertiary structure, but also quaternary structure where these forces contribute towards protein-to-protein interactions in the form of oligomerization. The formation of protein oligomers go on to have significant impact on the protein's stability and functionality [4]. For this reason, when studying the activity of a protein *in vitro*, attention must not be limited to a static image of the overall fold; there should be tantamount awareness of the forces that facilitate conformational changes and movements as well as those involved in oligomerization dynamics.

In Chapter Three, there will be a focused study of the mechanism of HMGR and will explore in detail the atomic-level level understanding of how HMGR works. This section, however, will preface those endeavors by focusing on the overall quaternary structural features of this enzyme and will use HMGRs from *Streptococcus pneumoniae* (SpHMGR) and *Delftia acidovorans* (DaHMGR). Using these two homologs as case studies, two primary areas that will be highlighted include 1) the overall three-dimensional dynamics, and 2) higher order oligomeric states. Both topics rely on the aforementioned forces that modulate secondary, tertiary, and

quaternary structure formation and movements and are discussed here. Also included, is a review of the technique of small-angle X-ray scattering (SAXS) used in this section to grasp the advantages and limitations of this approach in helping to unravel the fascinating structural considerations of class II HMGR.

### 2.2.1. Overall three-dimensional fold of class II HMGR

The first crystal structure of a class II HMGR was determined in 1995 and was a model of HMGR from *Pseudomonas mevalonii* (PmHMGR) determined at 3.0 angstroms [5]. PmHMGR, a 45 kDa protein, is like all HMGRs in that it is an obligate homodimer where two identical monomers come together to make up a dimer complex required for the reaction. In their structure, Lawrence et al. observed this obligate dimer in the asymmetric unit noting a large monomer-to-monomer interface of which a substantial portion is solvent-inaccessible [5]. This early model of PmHMGR did not include the last 50 residues of the enzyme and it was believed that this region was disordered in the crystal lattice [5]. Subsequently, PmHMGR structures at higher resolution, revealed novel locations of these last 50 residues and was later termed the C-terminal flap domain due to its perceived mobility during the course of the reaction[6].

The formation of the obligate homodimer includes two helices from each monomer, that contribute towards a familiar dimerization motif made up of an antiparallel four-helical bundle that constitute the core of the dimerization interface [5]. Dimerization allows the coming together of the active site where the substrate of one monomer reacts with the cofactor of the adjacent monomer. Figure 1 aids in highlighting the key overall structural features pertinent to this chapter.



Similarly observed in later structures of HMGR from *Streptococcus pneumoniae* (SpHMGR), the first ~70 residues of each monomer contain an interlocking motif that intertwine at residues 38-58 (SpHMGR numbering) and form an interlocked  $\beta$ -sheet [7]. A conserved sequence, ENVIGX<sub>3</sub>I/LP where X represents variable residues, makes up an important part of this dimerization motif[8]. Figure 2 shows a close-up of this dimerization interface. The dynamics of how this interlocking conformation is achieved is not known. For example, as each monomer is translated, do they fold independently, then undergo conformational changes to interlock with another monomer? Or do two monomers fold simultaneously into a dimer? Regardless of the method of dimerization, the nature of this motif suggests that there are concerted features that ensure that the dimer form remains intact.

The reason for this may lie in the fact that this N-terminal interlocking region contains important substrate binding residues, Glu50 and Asn51, such that these residues in one monomer interact with HMG-CoA of the adjacent monomer [7]. This observation, that there is an extensive interlocking interface with residues implicated in catalysis, supports the notion that the integrity of dimerization is pertinent to enzyme activity, and that the dimer conformation is an obligate active form.

Each HMGR monomer can be broken up into three domains: a large, a small, and a C-terminal flap domain [9]. The small domain, residues ~110-215, is made up of three anti-parallel  $\beta$ -sheets that stack up against two helices in an arc-like fashion and is linked to the large domain by two  $\beta$ -strands. The C-terminal flap domain, which extends from the opposite end of the molecule as the small domain, is made up for a three helical bundle linked to the large domain by a flexible strand and is made up of the last ~50 residues.

The rest of the molecule constitutes the large domain and is made up of ten helices and two sets of anti-parallel  $\beta$ -sheets. As noted by Lawrence et al., the large domain has a unique fold where a long helix, greater than 20 residues, runs through the center of this domain and is located within a wall of three separate substructures made up of a combination of helices and  $\beta$ -sheets with has an “arm” made up of the first ~60 residues at the N-terminus that extends away [5]. Figure 3 shows one monomer of SpHMGR as well as a corresponding two-dimensional topological outline of the secondary structural features with corresponding residue numbers generated by PDBSum [10].

When the monomers come together, without the C-terminal flap domain, they form a dimer that was initially described as having a “T” shape [5]. The large domains of each monomer primarily make up the central vertical component of the T whereas the small domains extend to either ends of the horizontal component. The C-terminal domain has been located in several distinct positions, as will be discussed in Chapter Three, and can span a wide area on the left or right of the overall T shape. In total, the overall structure of the homodimer contains both novel and previously observed structural motifs, an elaborate interlocking dimerization feature, and a flexible domain, all of which make HMGR a fascinating enzyme to study from a structural perspective. In addition to the dimerization motif discussed above, which points to the necessity of the obligate homodimer, evidence also points towards higher order oligomerization in some HMGRs and will be discussed next.

### 2.2.2. Higher order oligomerization

Since the early dissociation experiments conducted by Theodor Svedberg, which revealed that some proteins exist as multimeric complexes, extensive research has been done to assess not

only the role of oligomerization but also the forces that promote the coming together of protein complexes as well as the evolutionary mechanisms that have produced its roles [11] [12]. Protein-protein interactions rely on hydrogen bonding, salt bridges, and most notably, hydrophobicity, and in the context of oligomerization, can form either hetero or homo oligomers if either different or identical subunits come together, respectively [11], [13]. A significant amount of proteins form homo oligomers and this phenomenon has physiological implications on enzyme activity and stability but also on broader processes in the cell including pathway regulation, gene expression, and signaling [12]. HMGR oligomerization presents an interesting case study on this topic since there appears to be different oligomeric preferences among HMGR homologs.

#### 2.2.2.1. Class I HMGR

As pointed out by Istvan et al., early class I HMGR studies showed that mammalian HMGR, specifically rat HMGR, had a molecular weight of 50 kDa on a denaturing gel but 200 kDa in a size exclusion experiment [8]. This offered initial hints that the protein might exist as a multimer. This work, done by Brown et al., showed that at low temperatures the enzyme was inactive but with the addition of high amounts of salt, at 4 M, the high ionic environment was able to prevent the loss of activity [8], [14]. With this observation, they reasoned that the enzyme was a multimer held together by hydrophobic forces that were disrupted at lower temperatures but stabilized by electrostatic interactions with high salt [8], [14]. As previously mentioned, hydrophobicity plays a key role in protein folding where non-polar amino acids prefer being buried in the protein and polar amino acids being exposed to the solvent, and temperature and ionic state have different effects on this process[14].

Later work revealed that mammalian HMGR exists as a tetramer, as shown in Figure 4A. Armed with the crystal structure of the catalytic portion of human HMGR (hHMGR), it was observed that the aforementioned obligate dimers come together to form the tetramer and the authors propose that this tetramer formation was more stable than its dimer form [8]. In their crystal structure, Istvan et al. show that hHMGR forms tetramers with  $D_2$  symmetry, and note how the monomers from each dimer, extensively wind around each other as shown in Figure 4B, forming the active site that exists at the interface between two monomers similarly observed in PmHMGR above [8].

In its tetrameric form, about half of the surface is total solvent-inaccessible of which ~40% is involved in dimer formation and the rest involved in tetramer formation [8]. This extensive interface, observed in the crystal structure, showing a large contact area between the dimers, further substantiates the tetrameric form found in solution which was corroborated by analytical ultracentrifugation experiments[8]. In the case of hHMGR, the soluble catalytic portion is part of a larger assembly that includes a membrane region. It is believed that the soluble portion may facilitate tetramerization at the membrane region, and that dissociation of the tetramer membrane region exposes HMGR to protease activity leading to inactivation of the enzyme, thus serving as a mechanism of control[8]. The importance of this characteristics is that oligomerization is implicated in modulating the function of this enzyme, further emphasizing the significance of understanding this phenomenon as it relates to HMGRs in other organisms.

#### 2.2.2.2. Class II HMGR

In their first crystal structure of a class II HMGR, PmHMGR, Lawrence et al. observed that the dimer in the crystal structure was able to form a hexameric assembly made up of three

dimers formed around a threefold crystallographic axis [5]. However, it remained unclear whether this was a crystallographic artifact or whether a hexamer formation existed in solution. Subsequently, Rogers et al., performed gel filtration and sedimentation experiments and concluded that the active form in solution was indeed a hexamer. This hexameric form was later also observed in the crystal structure HMGR from *Burkholderia cenocepacia* (BcHMGR) [15]. It is worth noting that both PmHMGR and BcHMGR are NAD(H)-preferring HMGRs. HMGR from *Bordetella petrii* (BpHMGR) has not been fully characterized, but initial studies suggest that it is NAD(H)-preferring [16]. The Kung lab has solved structures of BpHMGR with various ligands bound (Raining Huo and Yan Kung, unpublished) which crystalizes an entire intact hexamer in the asymmetric unit.

Interestingly, using size exclusion chromatography, BcHMGR was shown to have potentially two other oligomeric states, a nonamer and a larger unknown assembly, such that the equilibrium between these species shifted based on the pH of its buffer and enzyme concentration [15]. In the case of hHMGR, formation of the tetramer buried about half of the surface area whereas in the hexameric assembly, Peacock et al point out that the hexamer buries a significantly smaller surface area and is more hydrophilic [15]. Their analysis of the hexameric interface revealed a “cylinder resting in the palm of a hand” configuration where a helix from one dimer rests across a four-stranded  $\beta$ -sheet on the adjacent dimer, and Peacock et. al point out that, in the case of BcHMGR, only a single salt bridge and two protein-protein hydrogen bonds hold this interface together [15]. Another unique finding in this study was that there exists a hexamer-to-hexamer interface, involving residues 8-17, containing an Arg11 that forms interactions with coenzyme A (CoA) [15]. When they computationally removed CoA and performed molecular dynamic simulations, they observed that Arg11 no longer interacts with CoA and instead with

Asp237, suggesting a dissociation mechanism of the hexamer-to-hexamer interface. The authors go on to propose that BcHMGR functionality can be described using the morpheein model of allostery where its catalytic activity is dynamically controlled by distinct oligomeric protein conformations[15], [17].

In 2018, the Kung Lab solved two structures of the NADPH-preferring HMGR from *Streptococcus pneumoniae* (SpHMGR), one with HMG-CoA bound (PDB 5WPK) and one with NADPH (PDB 5WPJ)[7]. The NADPH-bound structure formed two dimers in the asymmetric unit whereas the HMG-CoA-bound structure formed a dimer. In addition, two other apo structures of SpHMGR with no associated publication (PDB 3QAU and 3QAE) contain a single monomer in the asymmetric unit but crystallographic symmetry can create the dimeric assembly. This would suggest that SpHMGR exists as a dimer in the crystal lattice, but its solution state remains uncharacterized. The Kung Lab also recently solved the structure of HMGR from *Enterococcus faecalis* (EfHMGR) with no ligands bound and crystallized with a single monomer in the asymmetric unit (Jingyi Ren and Yan Kung, unpublished) that is also able to create the dimer using crystallographic symmetry. In a study of *Listeria monocytogenes* HMGR (LmHMGR), an HMGR that can use both NADP(H) and NAD(H) but has a preference for NADP(H), size exclusion chromatography revealed a mass of 84.9 kDa suggesting that it exists as a dimer in solution[18]. Taken together, oligomerization of numerous HMGRs appears to vary and warrants further investigation into the contributing factors.

So far, this chapter has examined the overall three-dimensional fold of HMGR and discussed the important features that have informed our structural perceptions of this intriguing enzyme. Much of this work has relied on X-ray crystallography which provides static snapshots of the enzyme as it is contained in the crystal lattice. X-ray crystallography, while potentially

helpful in its ability to predict the solution state oligomerization of the protein, is not a solution-based study and therefore is inconclusive in its assessments of how the protein might truly exist in its native state. We subsequently explored the published evidence of HMGR oligomerization in a variety of organisms, however, there is limited solution state analysis done that more definitely describe HMGR oligomerization. As such, the goals of this section will begin to address these areas. To preface these goals, however, the following section describes an important technique used that will be utilized in this chapter, that will help inform the nature of the findings.

#### 2.2.2.3. Solution-state technique - SEC-SAXS

Small-Angle X-ray Scattering (SAXS) of proteins is a solution-based technique that can provide insight into the structure of the protein as it exists in a given solution. Unlike traditional X-ray crystallography that requires the production of crystals and yields a static snapshot of the protein in the crystal lattice, SAXS has an inherent advantage of providing information that may be more closer to the protein's native structure since does not need to be confined to a crystal lattice. [19]. SAXS cannot provide high resolution structural models to the same extent as X-ray crystallography. It is, however, more suitable for providing high-precision data with regards to size and shape [20]. For this reason, SAXS often complements X-ray crystallography when assessing the structure and function of macromolecules. In this section, we will explore the fundamentals of SAXS and detail the various data analytical techniques used to better understand the results presented later in this chapter.

A SAXS experimental setup looks very similar to X-ray crystallography in that a high intensity light source provides an incident beam targeted at the sample, causing a resultant diffraction, in the case of crystallography, or scattering, in the case of SAXS, pattern on a detector

downstream of the beam direction. In X-ray crystallography, the diffraction pattern, or spots, is the result of constructive and destructive interference of the waves as they pass through repeated units of the protein in the crystal lattice. In SAXS, the scattering waves constructively or destructively interfere but due to the rotational averaging of the protein in solution, the resulting scattering pattern observed on the detector is isotropic i.e., the scattering is radially symmetric that is more intense closer to the center. The pattern provides intensity ( $I$ ) as a function of the amplitude of the scattering vector,  $q$ , where  $q = \frac{4\pi \sin\Theta}{\lambda}$ ; the angle ( $<3^\circ$ ) at which scattering occurs is at  $2\Theta$  generated from the angle formed by extrapolating the incident beam and the scattered x-rays [20]. Figure 5 shows an outline of a traditional SAXS experiment.

In an experiment, the solvent scattering is subtracted from the scattering of the protein in the solvent and yields plot of the scattering intensity versus the distance from the beam center. This plot is a scattering profile that directly, and precisely, relates to the shape of the protein. It is important to consider factors that influence monodispersity of the sample that must be taken into consideration when interpreting data. If one has a single protein conformation versus this same protein that exists in multiple oligomeric states, the scattering profile will give be different. A size exclusion chromatograph (SEC) is sometimes equipped upstream of the scattering experiment such that protein species can be separately assessed.

Structural information that can be obtained from the initial plot of scattering intensity versus the magnitude of the scattering angle is not immediately intuitive and further analysis is required to utilize the breadth of insight that can be achieved using this technique. One challenge of SAXS is that it typically requires highly pure samples since scattering is very sensitive to all moieties within a sample, including impurities and aggregation [20]. By including a Size-



Exclusion Chromatography apparatus upstream, as well as the ability to customize the buffer conditions to ensure protein purity, some of these concerns can be alleviated. In sum, SAXS can be a powerful technique that provides a wealth of information on the solution state of proteins, without any prior knowledge of the structure required or the formation of crystals.

### 2.2.3. Goals of this section

Prior to the work presented here on DaHMGR, the crystal structures of only three HMGRs were reported – hHMGR, PmHMGR and SpHMGR – of which only the latter two are from class II. The structure of a third class II HMGR, BcHMGR, was later solved. Considering how ubiquitous and central this enzyme is across biology, additional structures of HMGR remain a high priority. As such, one goal involves solving the crystal structure of apo-DaHMGR, to analyze its overall architecture in the context of previously reported observations, and to ultimately add to the above list of structurally characterized class II HMGRs. However, a necessary pre-requisite to this goal is to have a feasible approach of isolating soluble DaHMGR protein with relatively high concentrations and purity. In the case of DaHMGR, this required customization and optimization of purification techniques and will be discussed in detail.

In the studies that do have crystal structures, there is limited complementary biophysical analysis available. A preliminary evaluation of the available data suggests that oligomerization among the various characterized HMGRs vary. Moreover, questions into how and why oligomerization influences activity remain. Accordingly, the second goal of this section is to utilize the aforementioned biophysical techniques to explore the oligomerization of class II HMGR using DaHMGR and SpHMGR.

Importantly, in the case of DaHMGR additional work to complement this crystal structure, using techniques such as SEC and SAXS, will be employed. Since the crystal structure of

SpHMGR already exists, the goal then is to further investigate the biophysical characteristics of this previously studied HMGR, using SEC. What are the oligomeric states of DaHMGR and SpHMGR? Do they present themselves in predominantly one oligomeric state or an equilibrium between multiple states? Are there patterns in the oligomerization between various HMGR homologs that are related to their structure and/or function? Considering that, in the case of hHMGR, oligomerization is a mode of regulation, greater insight into the mechanisms of oligomerization may provide the basis for understanding the physiological relevance and the driving forces behind them.

## **2.3. Results**

### **2.3.1. Protein Expression and Purification of DaHMGR and SpHMGR**

Successful overexpression of DaHMGR and SpHMGR was observed by evaluating the protein makeup of cells pre- and post-induction prior to purification. When purifying DaHMGR, during early attempts, significant protein precipitation was observed during fractionation. Fractions containing the highest amount of protein precipitated shortly after fractionation. Upon further examination, it was also observed that during the incubation of protein with the Ni<sup>2+</sup> resin prior to fractionation, precipitation was evident. It was hypothesized that protein precipitation was concentration dependent and required an amended purification protocol.

To further optimize DaHMGR purification, during fractionation, volumes of as low as 0.5 mL/fraction were collected to prevent batch precipitation of large amounts. This necessarily caused some fractions to form precipitation, but other fractions that had moderate amounts of protein did not precipitate. Subsequently, fractions were centrifuged at high speed and combined with other

fractions with similar concentrations. It was observed that DaHMGR prefers concentrations at ~8 mg/mL or lower. This modified procedure allows for fractionation with greater control over the elution profile which enables the ability to acquire soluble protein at feasible amounts. SpHMGR, in contrast, purified without the above modifications and Figure 6 shows representative purification gels suggesting relatively high homogeneity using the above DaHMGR and SpHMGR purification protocols.

Since DaHMGR precipitated at high concentrations, buffer exchange into the storage buffer was not trivial. Initially, using a centrifugal concentrator, protein precipitation was observed due to the concentration gradient generated when centrifuging the protein in the concentrator. It was reasoned that a traditional dialysis, with the protein in dialysis membrane rotating in storage buffer, would be more appropriate. However, significant precipitation was observed using this method too. This could possibly be due to local pockets of high concentration protein during dialysis overnight, or it could be attributed to DaHMGR's intrinsic instability over time. Thus, a desalting column was successfully used to perform buffer exchange as outlined in the methods.

### 2.3.2. Overall structure of DaHMGR

Only the overall structure of DaHMGR is discussed here since the structure of SpHMGR was previously reported and incorporated into the discussion. Apo DaHMGR crystallized with one monomer in the asymmetric unit. Using crystallographic symmetry, the adjacent monomer that forms the obligate homodimer can be generated. Electron density for residues from 3-376 were observed and modeled in. No electron density for the C-terminal domain was observed, which is made up of residues 378-429. The active site and cofactor-binding sites are unoccupied. The structures of apo DaHMGR align well with other DaHMGR structures (discussed in Chapter

Three) as well as with those of other class II HMGRs, with a root-mean-square deviation for C $\alpha$  atoms (rmsd) of 0.75–0.84 using the structure of cofactor-bound SpHMGR (PDB entry 5WPJ). Figure 7 shows an overlay of apo DaHMGR and structures of HMGRs from other class II HMGRs.

### 2.3.3. Size Exclusion Chromatography (SEC)

#### 2.3.3.1. SpHMGR SEC

The size of SpHMGR expressed above is 49.2 kDa as a monomer, and therefore a dimer would constitute 98.4 kDa. The elution profile of SpHMGR shows that a single peak at approximately ~76 mL corresponds with a dimer. As shown in Figure 8, an overlay of protein standards and the elution of SpHMGR, SpHMGR elutes between the 44 kDa and 158 kDa protein markers and more convincingly resembles the size of a dimer.

#### 2.3.3.2. DaHMGR SEC

The size of DaHMGR expressed above is 48.1 kDa as a monomer, and therefore a dimer and hexamer would be 96.2 and 288.6 kDa, respectively. The elution profile of DaHMGR reveals four peaks as shown in Figure 9. The largest peak, which eluted at ~48 mL corresponds with a size >670 kDa and likely represents aggregation. The peak at ~63 mL represents a peak that is markedly larger than 158 kDa that may represent a hexamer and the peak at ~76 mL corresponds with a dimer. A peak at 88 mL suggests the presence of a species between 17 and 44 kDa that likely represents a contaminant. Upon examination of the SDS-PAGE gels, Figure 6, there is indeed a band representing protein between 30-40 kDa. Thus, in its soluble state, DaHMGR likely exists in an equilibrium between hexamers and dimers.

#### 2.3.4. DaHMGR SEC-Small-angle X-ray scattering

The size exclusion chromatograph of the sample revealed that there were multiply overlapping peaks, but it was possible to deconvolute three species: aggregate, a larger species, and a smaller species. Data from the larger and smaller species were processed separately. It is worth noting that only two peaks were analyzed, which does not rule out the presence of other species in the sample which would include a dimer. Only the two largest peaks were feasible for SAXS analysis as determined by the collaborators from the Ando Lab, Cornell University who assisted us with data analysis.

Depicted in the various plots on Figure 10, namely the Guinier, Kratky, and Logarithmic plots, is the larger species analyzed with respect to a dimer, tetramer, and hexamer. In these plots, the data is computationally compared to a theoretical fit of each of these oligomers. The larger species does not appear to resemble a dimer. Of the two higher order oligomers chosen to compare this species with, tetramer versus hexamer, the larger species more resembles a hexamer than a tetramer.

As depicted in similar plots on Figure 11, the smaller species also does not resemble a dimer. However, upon examining whether this species resembles a hexamer or tetramer, we observed no significant differences. Since there is no evidence that a tetramer formation is a feasible oligomerization state of class II HMGRs, we computationally simulated a dimer: hexamer ratio of 40%:60%, respectively, and noted that this configuration fits the tetramer and hexamer formations comparably as shown in Figure 12. This suggests, that instead of the smaller species resembling a tetramer, it may exist in an equilibrium between dimer and hexamer. This agrees with the SEC profile of DaHMGR in Figure 9.

## 2.4. Discussion

HMGR, a central enzyme found across all kingdoms of life, continues to be a gratifying protein of interest owed to its fascinating structural and functional dynamics that promise fruitful inquiry. Intriguing characteristics include structural features such as the mobile C-terminal flap domain, the non-Rossmann dinucleotide binding domain, and the dimerization motif of the obligate dimer. But despite its prevalence in biology, there are only a handful of crystal structures available for class II HMGR. Moreover, there is a lack of complementary biophysical analysis of how the protein behaves in solution. Characterizing a protein requires a multidisciplinary approach and depending on the nature of the investigation, one technique might be preferred over another, often requiring additional evidence that corroborates the conclusions.

In this chapter, one goal was to optimize the isolation of DaHMGR protein for X-ray crystallography experiments. Subsequently, by acquiring the crystal structure of DaHMGR, the overall three-dimensional fold was investigated in the context of previously solved structures of class II HMGRs. However, as alluded to, X-ray crystallography is limited in its ability to provide insight into questions about certain structural features of this enzyme including its native-state oligomerization. Therefore, a second goal was to investigate the oligomerization of HMGR to understand how this protein might exist in its native environment. This goal relied on solution-state studies that were discussed earlier in this chapter. In this discussion, we now critically reflect on and synthesize the observations gained from the above two goals, with the hope of further characterizing the structural features of this intriguing enzyme.

During purification, DaHMGR was especially sensitive to concentration. Precipitation was evident as early as during the lysis of *E. coli* expression cells and persisted throughout purification steps. An additional screen of buffers suggested that high amounts of salt, NaCl, glycerol, and

reducing agent, TCEP, all aided in stabilizing soluble protein and these conditions were utilized throughout protein purification steps. Following fractionation of the protein during affinity chromatography, the protein was collected and directly applied to a desalting column for the purpose of buffer exchange. Following fractionation from the desalting column, protein was stored for use. This protocol prevented centrifugation steps that are typically used during buffer exchange or concentrating of the protein using a centrifugal filter. While the purification protocol benefited from the additives and modified procedures above, higher concentrations of protein caused precipitation whereas concentrations below  $\sim 8$  mg/mL were more stable. Purification of SpHMGR was done according to a previously reported protocol [7].

Following the successful isolation of DaHMGR, initial protein crystals for DaHMGR were further optimized and diffracted to relatively high resolution and produced an apo model. Unsurprisingly, this apo model had no density for the C-terminal domain (CTD) since the inherent flexibility of this domain makes it difficult to capture a single static conformation. In the context of catalysis, the observation that the CTD would be mobile without any ligands emphasizes that closure of the CTD is associated with ligand-binding. This phenomenon would allow ligands to be retained during the course of the reaction. There is, however, a crystal structure of SpHMGR (PDB 3QAU) that is both apo and has a resolved CTD, however, the CTD adopts an open conformation. This would suggest that in apo form, a flexible and open conformation predominates such that ligands are free to enter either cofactor or substrate sites. The role of the CTD is further investigated in Chapter Three.

The remaining overall structure of apo DaHMGR resemble the structures of other class II HMGRs. The conserved dimerization motif, ENVIGX<sub>3</sub>/LP, which is ENVFGSFELP in DaHMGR, indeed forms an intimate interlocking motif that aids in the formation of the

homodimeric interface. Interestingly, whereas other structures of HMGRs observe this region as a B-strand, apo DaHMGR does not form this secondary structure and instead appears as a long strand. Despite this observation, when overlaying this motif with the same region of PmHMGR, we observe that the side chains of the above residues align almost identically suggesting that the integrity of this motif is still maintained. Figure 13 shows an overlay of the dimerization motif of apo DaHMGR and apo PmHMGR.

Using crystallographic symmetry, a hexamer made up of monomers of apo DaHMGR can easily be generated. Figure 14 shows the hexamer, which forms a trimer of dimers, of apo DaHMGR comprising residues 3-376 in each monomer and excludes the CTD. The hexameric interface is made up of a helix of one monomer that stacks up against an anti-parallel beta sheet. This was later described by Peacock et al. as a “cylinder resting in the palm of a hand” configuration where a helix from one dimer rests across a four-stranded  $\beta$ -sheet on the adjacent dimer, as described earlier [15]. With the help of the crystal structure, these interactions can be parsed out, residue-by-residue, and as pointed out by Peacock et al. a total of two potential salt bridges and 8 hydrogen bond interactions exist at the hexameric interface of DaHMGR [15]. They point out that, in comparison to BcHMGR and PmHMGR hexameric structures, DaHMGR has the most interactions suggesting that its preference for the hexameric form is notably higher.

The following interactions at the hexameric interface of DaHMGR are worth noting. Glu308 from one monomer appears to interact with the backbone of His125 and Gly126 of the adjacent monomer. Gln133 of one monomer interacts with the neighboring monomer via the backbone of Thr105, and the sidechain of Gln104. Peacock et al. also suggest that two potential salt bridges exist between Asp123 of one monomer and Arg226 of the adjacent monomer as well



as Glu320 of one monomer and Arg128 of the adjacent monomer [15]. Figure 15 depicts these potential interactions at the hexameric interface of DaHMGR.

Using the crystal structure of DaHMGR determined from the present study, analyzed by Proteins, Interfaces, Structures and Assemblies (PISA) analysis, Peacock et al. point out that, in DaHMGR, the hexameric interface is the most stable assembly showing the hexamer as the most stable conformation ( $\Delta G_{\text{diss}} = 15.6$  kcal/mol). In contrast, PISA analysis of SpHMGR crystal structure (using PDB 5WPJ) suggests that SpHMGR exists as a dimer ( $\Delta G_{\text{diss}} = 83.9$  kcal/mol).

While the above analysis of oligomerization provides great insight into the features that promote these interactions and provides reasonable analysis of the above conclusions, X-ray crystallographic snapshots cannot conclusively say how the protein exists in its native state. Moreover, the crystal lattice may have captured one oligomeric state, but this does not provide evidence for whether this is the only oligomeric state observed in solution. Thus, solution state studies of DaHMGR and SpHMGR come into focus.

Using Size Exclusion Chromatography (SEC) to assess DaHMGR oligomerization, there are four notable peaks that correspond to aggregation, a likely hexamer, a dimer, and a contaminant. The sizes of the peaks suggest that the hexameric assembly predominates. This suggests that DaHMGR exists in equilibrium between a dimer and hexamer, where a hexamer is more abundant. When higher concentrations of DaHMGR are assessed, this equilibrium reveals an even greater abundance of hexamers than dimers in solution (Ashley Yang and Yan Kung, unpublished). Preliminary SAXS experiments, which assessed the solution state of DaHMGR also pointed to a hexamer formation. In these experiments, a SEC was fixed upstream of the SAXS experiment, separating the protein based on size prior to scattering experiments. Peaks that were

chosen to be analyzed included the predominant two peaks, and inspection of these peaks with Oligomer suggested that a hexamer is likely. The SAXS experiments do not rule out the presence of other species in solution but confirms the likelihood of the hexamer being in solution.

Taken together, the crystal structures, PISA analysis, SEC elution profiles, and preliminary SAXS experiments, all point towards DaHMGR existing in an equilibrium of multiple oligomeric states, with a hexamer being especially favored state at higher concentrations. Interestingly, as pointed out earlier, DaHMGR tends to precipitate at high concentrations as revealed through aggregation peaks in SEC experiments and during protein purification. The physiological relevance of this remains unknown. While the hexamer has been shown to be the most stable assembly, by PISA analysis, it is unknown why higher concentrations, which favor hexamer formation, leads to precipitation when the concentration goes above a certain threshold.

Regarding SpHMGR, the crystal structures determined here, PISA analysis, and SEC elution profiles, all point towards SpHMGR existing as a dimer. Upon close inspection, there is a marked difference in the types of residues that constitute the hexameric interface in DaHMGR when compared to the aligned residues in SpHMGR. As shown in Figure 16, if one aligns the adjacent monomers of the hexameric assembly with monomers of SpHMGR, there are several steric clashes observed. In Figure 16, residues such as Asn121, Lys123, Leu124, and Glu127 all appear to sterically clash with the adjacent monomer in the hexameric configuration. These observations provide structural evidence that the hexameric formation is hindered in SpHMGR.

More specifically, one could look at each of the interactions observed at the hexameric interface in DaHMGR outlined in Figure 15 and observe the lack of these attractive forces in SpHMGR. The following interactions, noted above, are observed at the hexameric interface in

DaHMGR: Gln133 interacts with the side chain of Gln104; Glu308 interacts with His125 and Gly126; Asp123 interacts with Arg226; Glu320 interacts with Arg128. In SpHMGR serine and threonine replace Gln133 and Gln104 found in DaHMGR. In DaHMGR the glutamine sidechains can interact each other, whereas serine and threonine found in SpHMGR cannot. In SpHMGR, asparagine and serine replace Asp123 and Arg226 found in DaHMGR. In the case of DaHMGR, Arg226's long side chain can extend closer to Asp123 to facilitate the sidechain interactions, whereas asparagine and serine in SpHMGR are  $>7 \text{ \AA}$  away which would prevent sidechain interactions. This is similarly observed in the interaction between Glu320 and Arg128 in DaHMGR which is replaced by threonine and asparagine, respectively, in SpHMGR. While it is possible that the sidechains of threonine and asparagine can interact, they are  $\sim 6 \text{ \AA}$  away from each other. Collectively, the lack of stabilizing interactions and the steric clashes that hinder the hexameric formation provide structural reasoning for why SpHMGR might not prefer the dimer formation and exists as a dimer.

Further examination of the hexameric interface provides insight into the types of HMGR homologs that prefer the hexameric formation with regards to cofactor specificity. As pointed out by Peacock et al. PmHMGR, DaHMGR and BcHMGR all contain the interactions between Glu320 and Arg128 that stabilizes the hexameric interface. Sequence alignment of this region, shown in Figure 20A, confirms that all the HMGRs whose oligomerization states have been shown to include a hexamer contain this glutamate residue which interacts with the corresponding arginine. These all appear to be NAD(H)-preferring HMGRs. In contrast, NADP(H)-preferring HMGRs, which have been shown to form dimers, do not have this glutamate and have shorter residues such as threonine and serine.

There is now mounting evidence that oligomerization is directly related to cofactor specificity. There appears to be a developing pattern where NAD(H)-preferring HMGRs, such as PmHMGR, BcHMGR, BpHMGR and DaHMGR, can form hexamers, whereas NADPH)-preferring HMGRS, such as EfHMGR and LmHMGR do not. Figure 20 shows a list of HMGRs, their cofactor preference, and their associated oligomeric states. In addition, a phylogenetic tree using the amino acid sequences of the above HMGRs generated using PhyML[22] reveals how the NAD(H)-preferring and NADP(H)-preferring HMGRs are more closely related to each other as revealed in Figure 20B. It appears that evolution has produced lineages that manifest itself through cofactor preference and oligomerization in class II HMGRs. Importantly, the data obtained here strengthens this paradigm and offers insight into the structural mechanisms that promote such oligomeric dynamics.

In sum, this chapter provides structural and biophysical characterization of DaHMGR and SpHMGR that sheds light on the overall structural features of class II HMGR, using X-ray crystallography, as well as insight into the way the protein behaves in its native environment using solution state studies. In doing so, the results from this chapter add to key areas when characterizing HMGRs by extending existing knowledge about the structures of these enzymes as well as taking steps towards areas that deserve more attention. With a greater awareness of the overall structural features of this enzyme, a closer look at features that govern the mechanism is presented in the next chapter.

## **2.5. Future work**

### **2.5.1. Stability of DaHMGR**

As pointed out, purification of DaHMGR is difficult due to its tendency to precipitate at higher concentrations. While there is the possibility that by customizing the buffer constituents during purification that the protein may be more stable, it is also likely that this phenomenon is inherent to the way the protein behaves. It would be worthwhile to investigate this further by attempting to manipulate the purification protocol as well as observe how purified protein behaves in different environments. Is it just a matter of including the right stabilizing reagents to prevent precipitation, or is there a fundamental feature of DaHMGR, which is related to its concentration, that is more influential?

It would also be helpful to determine if protein, which has precipitated due to high concentrations, can be resolubilized, and to determine whether the activity of this reconstituted protein is affected. If the activity can be restored effectively, purification of DaHMGR and downstream steps, including buffer exchange and concentrating of protein, can be more readily and conveniently achieved. If, on the other hand, the activity of precipitated protein is irreversibly diminished, then this would also inform interpretation of activity assays with regards to specific activity. One must ensure that the amount of protein, which is crucial when calculating specific activity, is accurately assessed with regards to how much soluble versus insoluble protein there is present in a sample.

#### 2.5.2. Oligomerization and activity

Since DaHMGR can exist in an equilibrium of hexamers and dimers, further analysis can be done to investigate the factors that contribute towards the various oligomerization states which could include investigating protein concentration, and the salt and pH of protein buffers. In addition, further work can also be done to determine what, if any, influence oligomerization has

on activity. Does a dimer perform differently than a hexamer? Are there cooperativity dynamics with respect to binding of ligands and on overall activity that relate to the hexamer? To this end, novel structures of NAD(H)- and NADP(H)-preferring class II HMGR homologs might provide further insight into oligomerization dynamics. Ultimately, these questions would provide insight into the physiological relevance of oligomerization in class II HMGRs.

## **2.6. Materials and methods**

### 2.6.1. Isolating DaHMGR and SpHMGR

#### 2.6.1.1. Genes

The DaHMGR and SpHMGR genes were inserted into pSKB3, a modified pET28b plasmid which encodes an N-terminal TEV protease-cleavable hexahistidine tag and a kanamycin resistance cassette, by previous members of the Kung Lab. Figure 17 details the DNA sequence including the promoter and terminator as well as the restriction digestion sequences of both genes. The plasmids were then transformed into chemically competent BL21(DE3)\* *E. coli* by adding 1  $\mu$ L of DNA into 50  $\mu$ L of cells supplemented with 10  $\mu$ L of 5xKCM (0.5 M KCl, 0.15 M CaCl<sub>2</sub>, 0.25 M MgCl<sub>2</sub>) and water up to a total volume of 100  $\mu$ L. The cell mixture was incubated on ice for 30 minutes, heat-shocked at 42°C for 90 seconds, and cooled on ice for 2 minutes, after which the cells were allowed to recover and proliferate at 37°C with shaking at 200 rpm for 1 hour. Then, 50  $\mu$ L of the cell mixture was spread onto lysogeny broth (LB) agar plates containing kanamycin (50  $\mu$ g/mL), allowing the cells to grow at 37°C overnight, yielding single colonies of cells that contained the plasmid with the gene encoding the proteins of interest. The target protein sequences that are encoded by these genes are shown in Figure 18.

### 2.6.1.2. Protein Expression

Both DaHMGR and SpHMGR were expressed similarly as outlined here. A single colony from the transformation was used to generate a starter culture containing LB with kanamycin (50  $\mu\text{g}/\text{mL}$ ), which was grown at 37°C overnight with shaking at 200 rpm. Then, 1 L of LB supplemented with kanamycin (50  $\mu\text{g}/\text{mL}$ ) was inoculated with 1% (10 mL) of the starter culture, grown at 37°C at 200 rpm until cells reached  $\text{OD}_{600} \sim 0.6-0.8$ . Protein expression was induced using isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM at 16-20°C at 200 rpm for 16-18 hours. The expression efficiency was assessed by PAGE. Cells were harvested by centrifugation at  $5,000 \times g$  for 12 min, flash frozen in liquid nitrogen, and stored at -80°C.

### 2.6.1.3. Lysis Buffer Screen for DaHMGR

Prior attempts within the Kung Lab at purifying DaHMGR proved unsuccessful due to precipitation during the purification process. To determine a suitable lysis buffer, and therefore a purification buffer system, several lysis buffers were assessed based on DaHMGR's abundance in the soluble layer (aqueous layer) after cell lysis and clarification as determined by SDS-PAGE. The list of buffers used is shown below in Figure 19. The screen was performed by suspending  $\sim 0.2$  mg cell pellet in 2 mL lysis buffer, sonicating at 25% amplitude using 3 cycles of the following method: 5 second pulse, rest on ice for 2 minutes. The lysate was clarified by centrifugation at  $14,000 \times g$  for 20 min at 4°C and the supernatant and pellet layers were run on an SDS-PAGE gel to determine whether DaHMGR precipitated into pellet or remained soluble in the aqueous layer. While DaHMGR showed similar solubility in many of the buffers, a Tris buffer close to physiological pH was determined to be the most suitable. This optimized buffer, 50 mM

Tris pH 7.8, 400 mM NaCl, 1 mM TCEP, and 10% glycerol, was the basis for all subsequent protein purification are shown in Figure 6.

#### 2.6.1.4. Protein Purification

##### 2.6.1.4.1. DaHMGR

Cells were thawed and resuspended in lysis buffer supplemented with lysozyme, 1 U/mL Pierce Universal Nuclease and 1.0 mM phenylmethanesulfonylfluoride (PMSF). To lyse the cells, sonication was performed on ice at 40% amplitude using four cycles of the following method: 3 second pulse, 5 second rest, 1.5-minute total pulse time, allowing 2-3 minutes of rest between cycles to prevent the sample from overheating. Clarification of the lysate was achieved by centrifugation at  $\sim 37,000 \times g$  for 30 min at 4°C to separate any insoluble protein and cell debris. A column containing 2 mL Ni-nitrilotriacetic acid (Ni-NTA) resin per liter of cell culture was pre-equilibrated with 3 column volumes (CV) of lysis buffer. The clarified supernatant was incubated with the resin at 4°C with gentle rotating for 5-7 minutes (longer incubation times caused protein precipitation), after which 10 CV of lysis buffer, and 10 CV of wash buffer was passed through the column. DaHMGR was eluted using elution buffer in  $\sim 1$  mL fractions. Fractions with highest purity and concentration assessed by SDS-PAGE were collected in preparation for buffer exchange and storage.

Initial attempts using various buffer exchange methods, using dialysis, or using a centrifugal concentrator, caused precipitation of the protein. Precipitation was also observed when protein concentration was attempted using a centrifugal concentrator. To overcome these challenges, buffer exchange was performed using a desalting column collecting the fractions with the highest concentration for further use without any subsequent concentration steps. A PD-10 (Ge



Healthcare) column was pre-equilibrating with storage buffer and the purified protein was then applied to the column and eluted with storage buffer in ~ 1 mL fractions, which were assessed by SDS-PAGE, and fractions with the highest concentration and purity were pooled to obtain final concentrations of 5-8.5 mg/mL as determined by the  $A_{280}$  using a calculated extinction coefficient of  $21,500 \text{ M}^{-1} \text{ cm}^{-1}$  and a molecular weight of 48.1 kDa. Aliquots of DaHMGR were flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

#### 2.6.1.4.2. SpHMGR

The purification for SpHMGR followed the published protocol by Miller et al [7] and followed a similar protocol as for DaHMGR above with the following modifications: the buffers can be found on Figure 6; fractionation when eluting from the Nickel column was done in ~ 3 mL volumes, and the fractions with highest concentration and purity were combined; buffer exchange was performed using a centrifugal filter. Final concentrations of protein, as determined by the  $A_{280}$  using  $38,900 \text{ M}^{-1} \text{ cm}^{-1}$  and a molecular weight of 49.2 kDa, was ~10 mg/mL. Aliquots of SpHMGR were flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

### 2.6.2. Crystallization

#### 2.6.2.1. DaHMGR

The Materials and Methods section in Chapter Three provides greater depth into the initial high throughput crystallization experiments that yielded crystal hits. These initial crystal hits of DaHMGR, with NADH bound, contained citrate in the substrate binding site due to the crystallization buffer being made up of 100 mM citrate. Following the acquisition of DaHMGR-bound NADH and citrate, we sought to acquire additional structures of DaHMGR without citrate

bound. This was done by replacing citrate in the crystallization condition with other buffers, adjusting pH, and including other precipitating reagents. By modifying the crystallization buffer to Bis-Tris pH 5.5-6.5 and adding polyethylene glycol 3350 ~18-19%, we were able to acquire crystals of apo DaHMGR that were suitable for diffraction experiments. This apo structure was optimized to contain 100 mM Bis-Tris pH 5.5, 600 lithium sulfate, and 18% PEG 3350. For diffraction experiments, the crystals were soaked in a cryoprotectant solution that contained a 1:1 volume ratio of the crystallization solution and a solution containing 16% ethylene glycol, 16% glycerol, 18% sucrose, and 4% glucose before being flash-cooled in liquid nitrogen.

#### 2.6.2.2. X-ray Data Collection, structure determination, and refinement

X-ray diffraction data were collected at the Advanced Photon Source (APS). Data was indexed, merged, and scaled using MOSFLM in space group *P321* with a single molecular in the asymmetric unit. Since the NADH/citrate structure was determined first, and that the apo structure crystallized with a different unit cell and space group, *P6<sub>5</sub>22* versus *P321* respectively, molecular replacement was used to solve the apo structure with the NADH/citrate structure as the search model in Phaser in the PHENIX suite. Iterative rounds of reciprocal space refinement were performed, with real space refinement done in Coot. The final data collection and refinement statistics of this structure, at 1.93 Å resolution, is presented in Table 1.

#### 2.6.3. Small-angle X-ray Scattering

##### 2.6.3.1. Sample preparation

Samples of DaHMGR were transported on ice to the Cornell High Energy Synchrotron Source (CHESS). Precipitation was observed during transit, most likely because the protein was

no longer frozen for an extended time. Prior to injection, samples were centrifuged at high speed and resulted in a protein concentration of ~4.7 mg/mL as assessed by  $A_{280}$ .

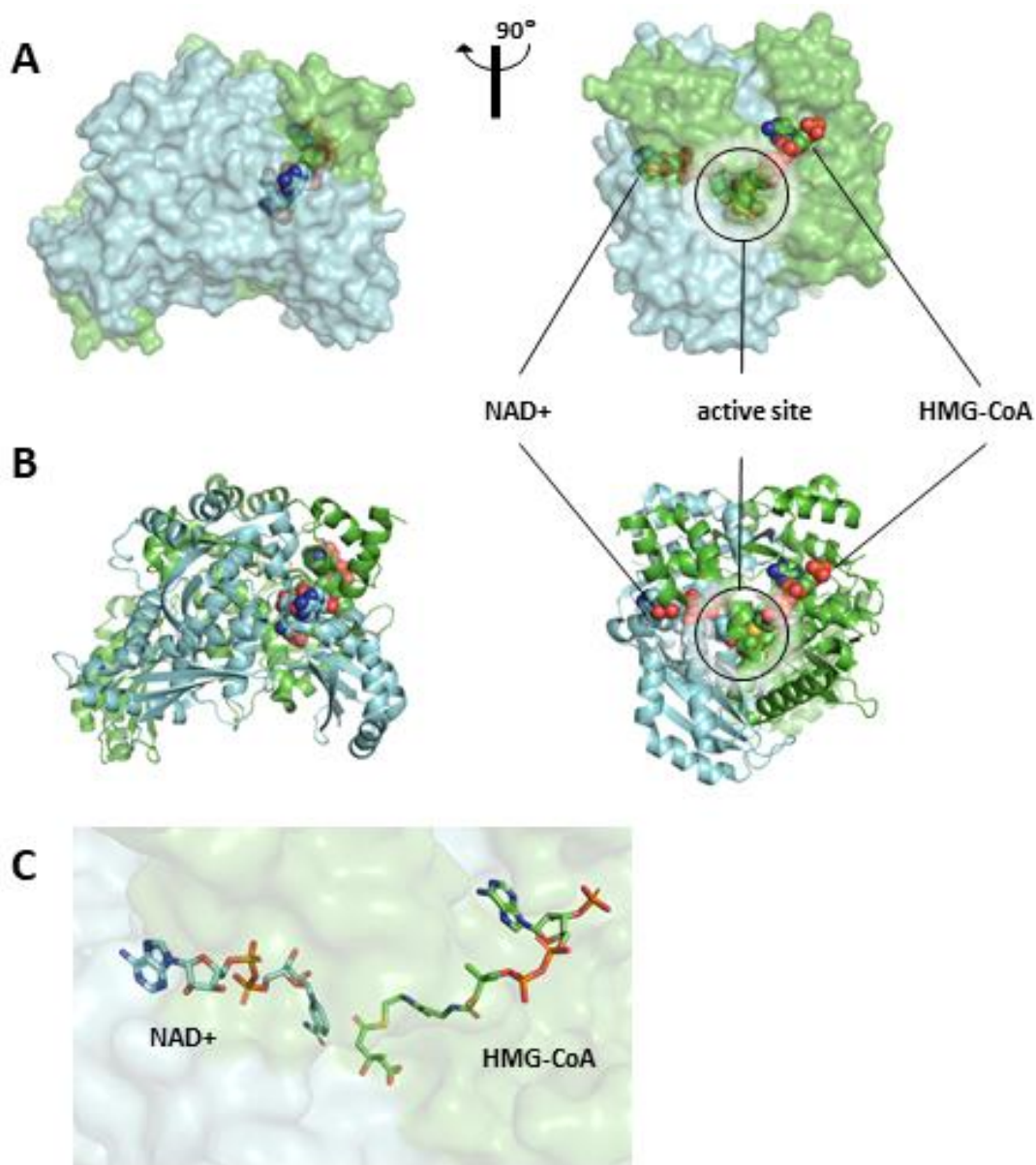
#### 2.6.3.2. SEC-SAXS data collection

Samples were run on a SAXS beamline at Cornell High Energy Synchrotron Source equipped with a SEC upstream. The SEC column was a Superdex 200, 5/150 (2.945 mL column volume). DaHMGR storage buffer, 50 mM Tris pH 7.8, 400 mM NaCl, 1 mM TCEP, and 10% glycerol, was used as the running buffer with a flow rate of 0.17 mL/min. Data collection was performed as follows: 1,000 frames, 2 sec exposure (to increase signal). Data was analyzed using Primus, Oligomer, and FOXS with the help of members of the Nozomi Ando Lab, Cornell University.

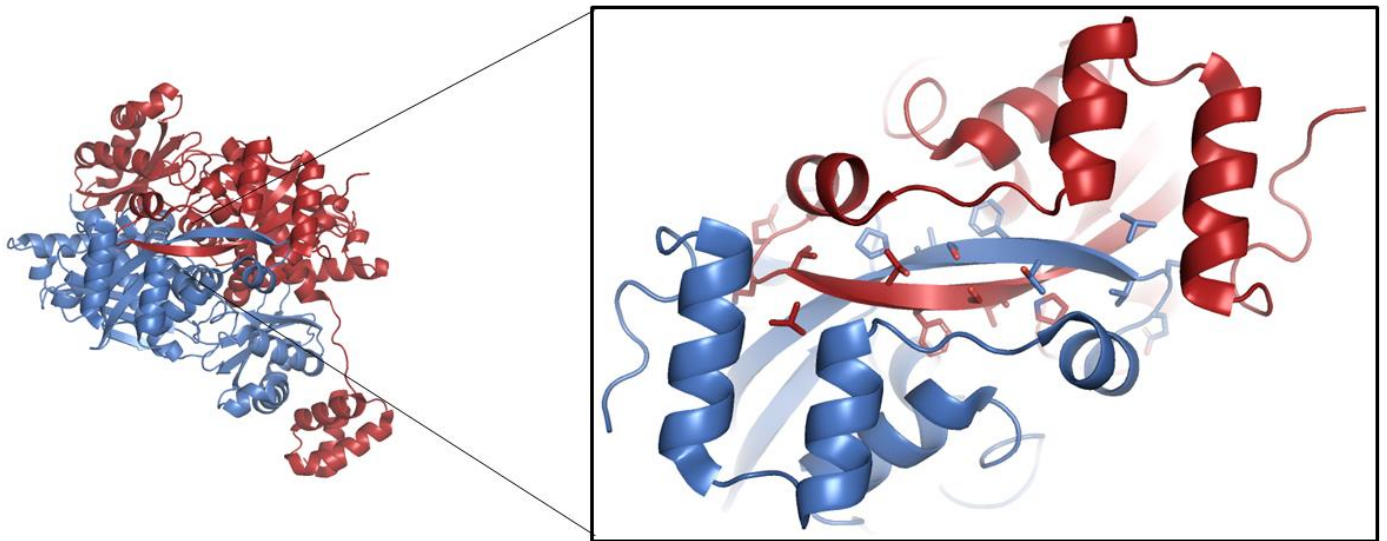
#### 2.6.4. Size Exclusion Chromatography

The SEC column was a HiLoad 16/600 Superdex 200 prep grade column (GE Healthcare Life Sciences) equilibrated with storage buffer (DaHMGR = 400 mM Tris pH 7.8, 400 mM NaCl, 1 mM TCEP, 10% glycerol, SpHMGR = 50 mM Tris (pH 7.8), 200 mM NaCl, and 10% glycerol) using an Akta Pure chromatography system (GE Healthcare Life Sciences). Injection volumes and concentrations were as follows: DaHMGR, ~3 mg/mL, 0.75 mL injection volume; SpHMGR at ~10 mg/mL, 0.75 mL injection volume. Protein standards were run in buffers resembling the storage buffer utilized for each protein. Elution was monitored by absorbance at  $A_{280}$ .

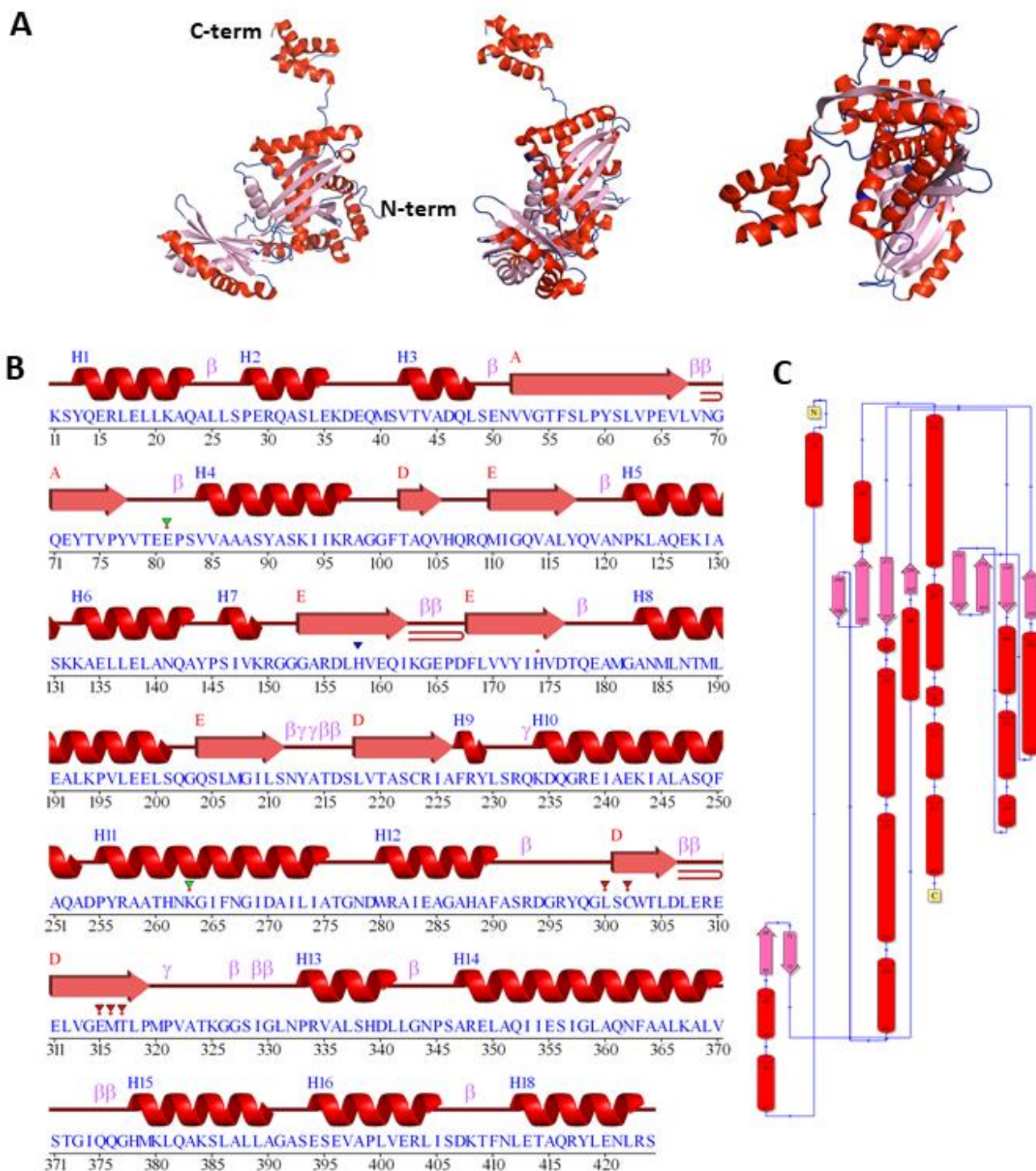
## 2.7. Tables and figures



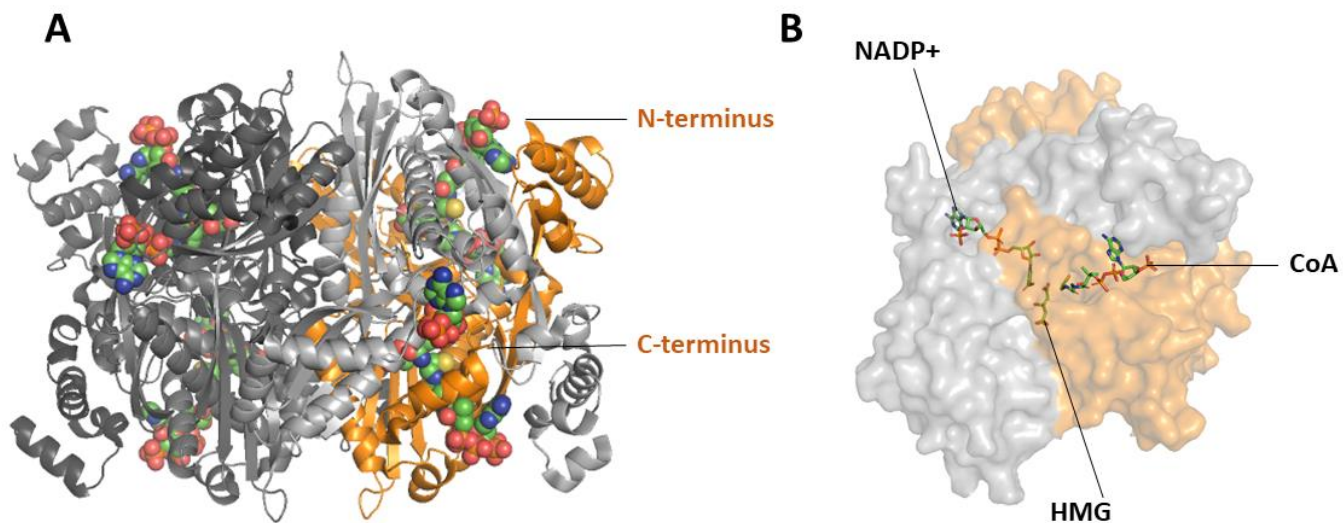
**Figure 1. PmHMGR crystal structure in a ternary complex.** PmHMGR dimer (PDB 1QAX) bound to NAD<sup>+</sup> and HMG-CoA with one monomer colored green and the other blue highlighting the active site and the orientation of the ligands. **A)** Two views in surface representation with ligands represented as spheres with C atoms of HMG-CoA colored green, the N atoms colored blue, the O atoms colored red, the P atoms colored orange, and the S atoms colored yellow. **B)** Two views in cartoon representation with ligands represented as balls. **C)** Zoom in on the active site with ligands represented as sticks in blue (NAD<sup>+</sup>) and green (HMG-CoA) and specific atoms colored similarly to the spheres.



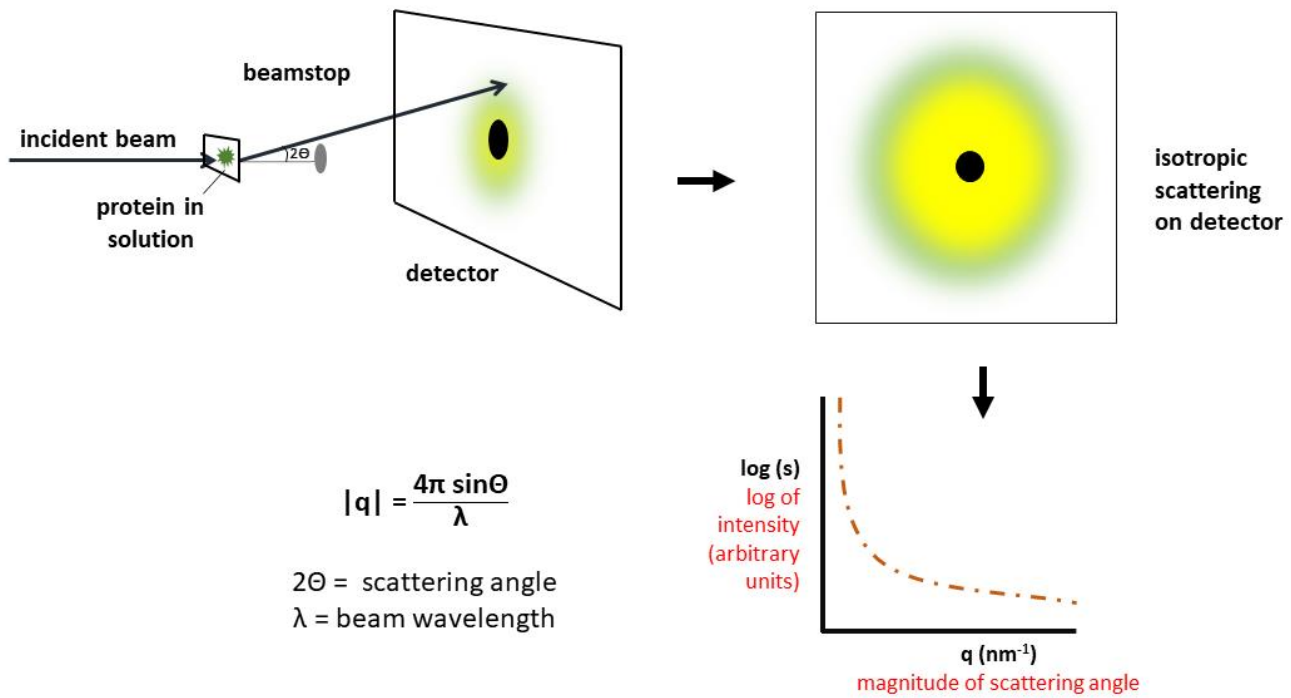
**Figure 2. Dimerization motif of SpHMGR.** SpHMGR (PDB 5WPJ) homodimer with monomers colors blue and red. A close-up of the dimerization interface is highlighted, with the conserved ENVIGX<sub>3</sub>ILP sequence in sticks colored by chain and the N atoms colored blue, the O atoms colored red.



**Figure 3. Overview of the structure of one monomer of SpHMGR.** A) Three different views of one monomer of the crystal structure of SpHMGR (PDB 5WPJ chain B) shown in cartoon representation. Secondary structures are colored similarly to the topological outline shown in C. B) Schematic generated by PDBSum that shows the residue numbers and predicted secondary structures of SpHMGR. C) Topological depiction of SpHMGR generated by PDBSum .

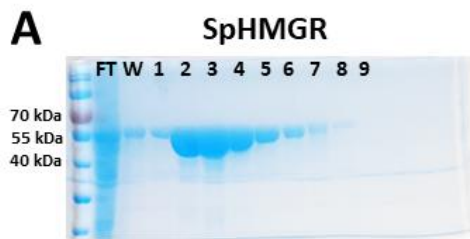


**Figure 4. Crystal structure of human HMGR.** **A)** The tetrameric conformation, made up of two dimers, with three monomers colored shades of gray and one monomer in orange. The termini of the orange monomer are noted. In balls are HMG, NADP<sup>+</sup> and CoA bound. **B)** A dimer, with one monomer colored gray and another orange, showing the interface as well as the location of the ligands depicted as sticks.

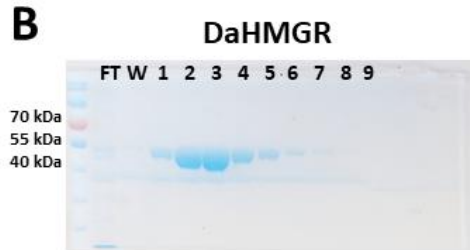


**Figure 5. SAXS overview.** An outline of SAXS from experimental setup to data collection.



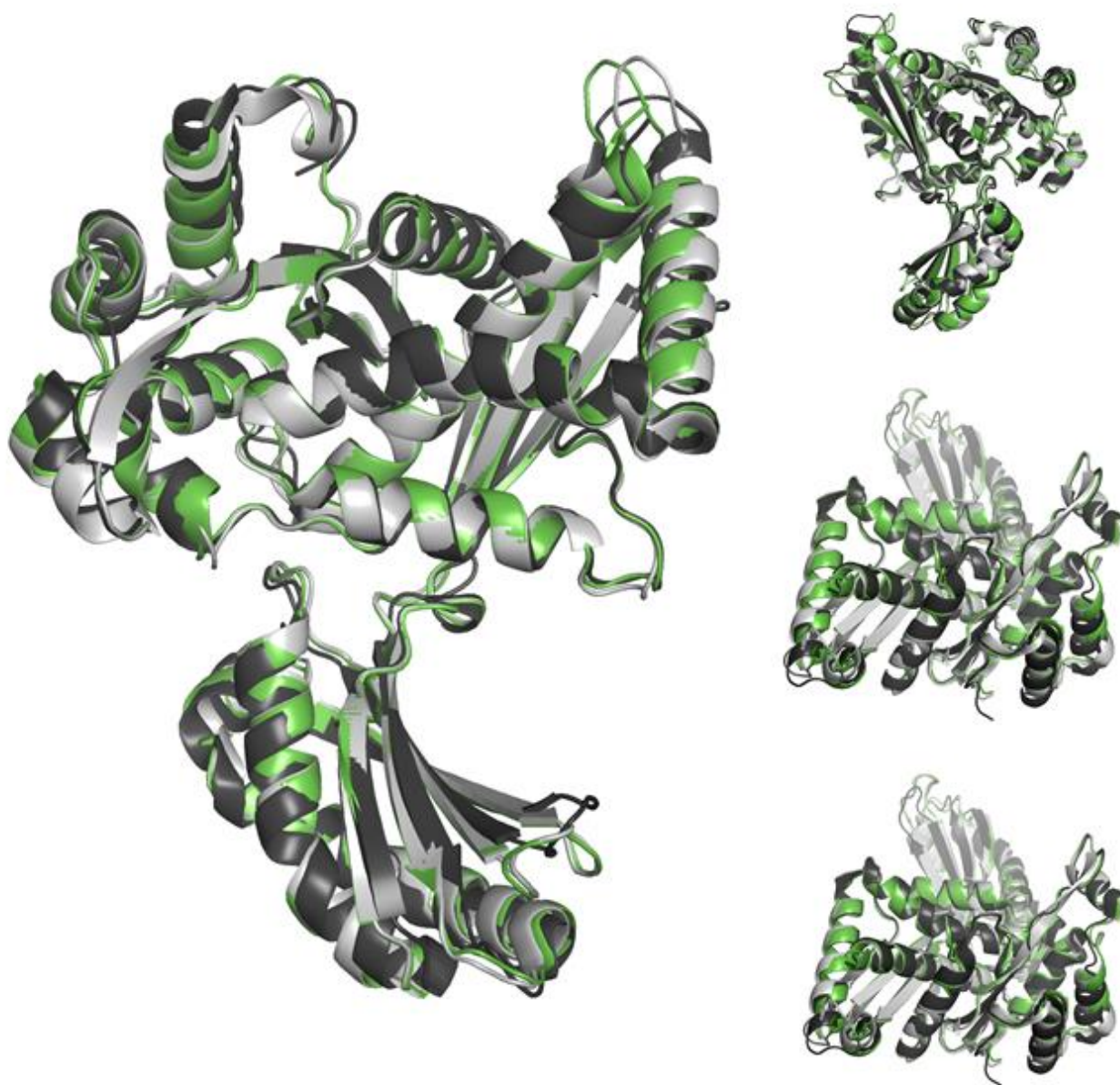


	Lysis (mM)	Wash	Elution (mM)	Storage (mM)
Tris (pH 7.8)	50	10% elution	50	50
NaCl	200	buffer in lysis	200	200
Imidazole (pH 7.8)	10	buffer	250	-
Glycerol	10%		10%	10%

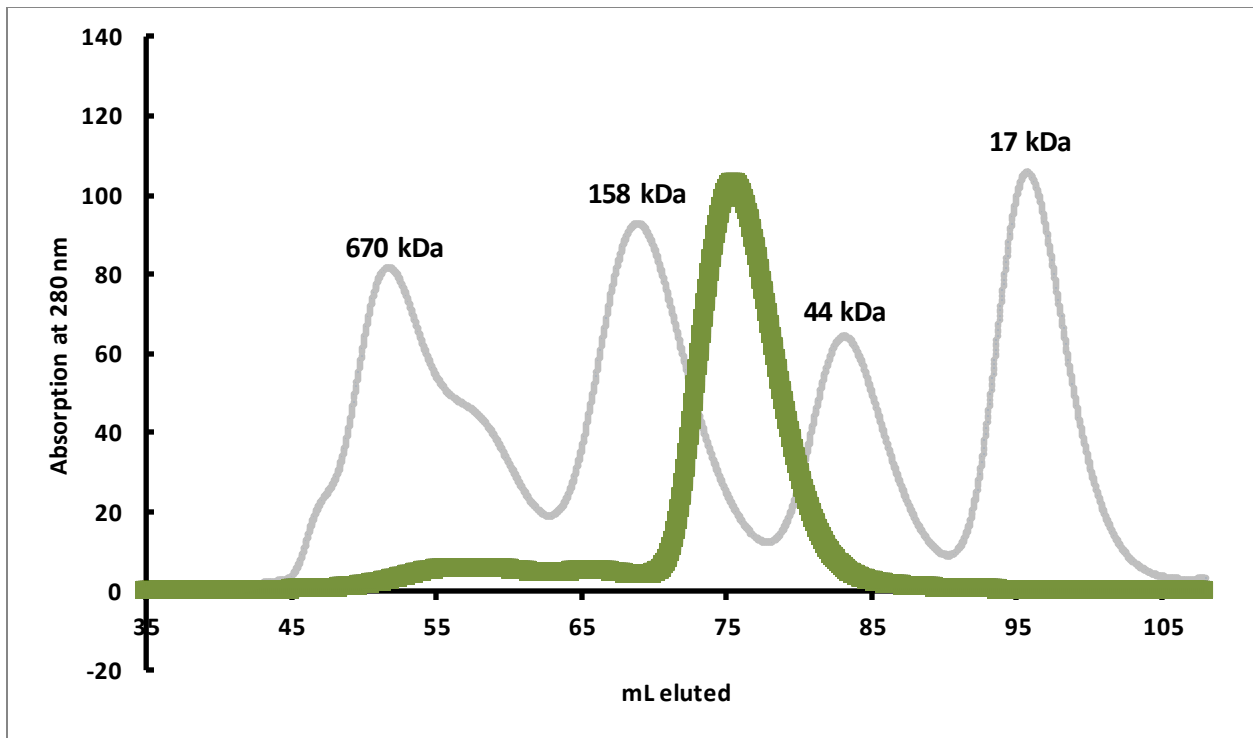


	Lysis (mM)	Wash	Elution (mM)	Storage (mM)
Tris (pH 7.8)	50	10% elution	50	50
NaCl	400	buffer in lysis	400	400
Imidazole (pH 7.8)	10	buffer	250	-
TCEP	1 mM		1 mM	1 mM
Glycerol	10%		10%	10%

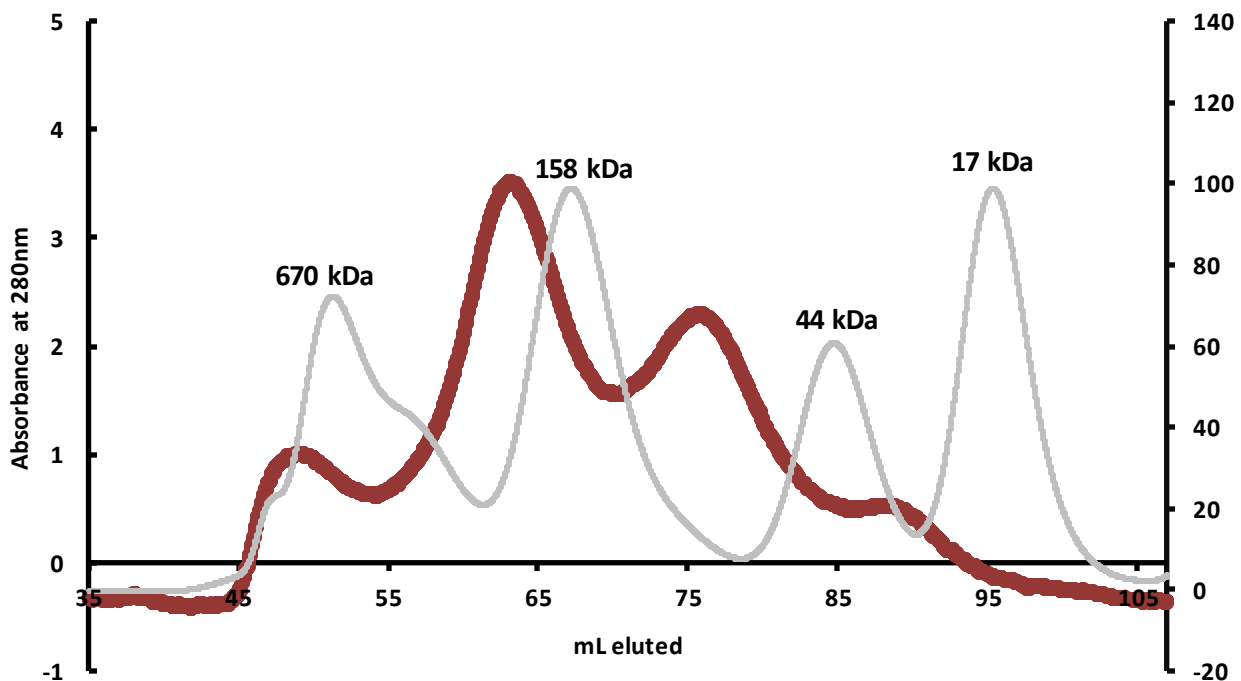
**Figure 6. Purification of DaHMGR and SpHMGR.** On the left are SDS-PAGE gels representing the following fractions collected during purification using Ni-NTA affinity chromatography: FT = flowthrough, W = wash, 1-9=fractions. On the right outlines the makeup of the various buffers used throughout Ni-NTA purification and storage. Purification with respect to **A)** SpHMGR and **B)** DaHMGR.



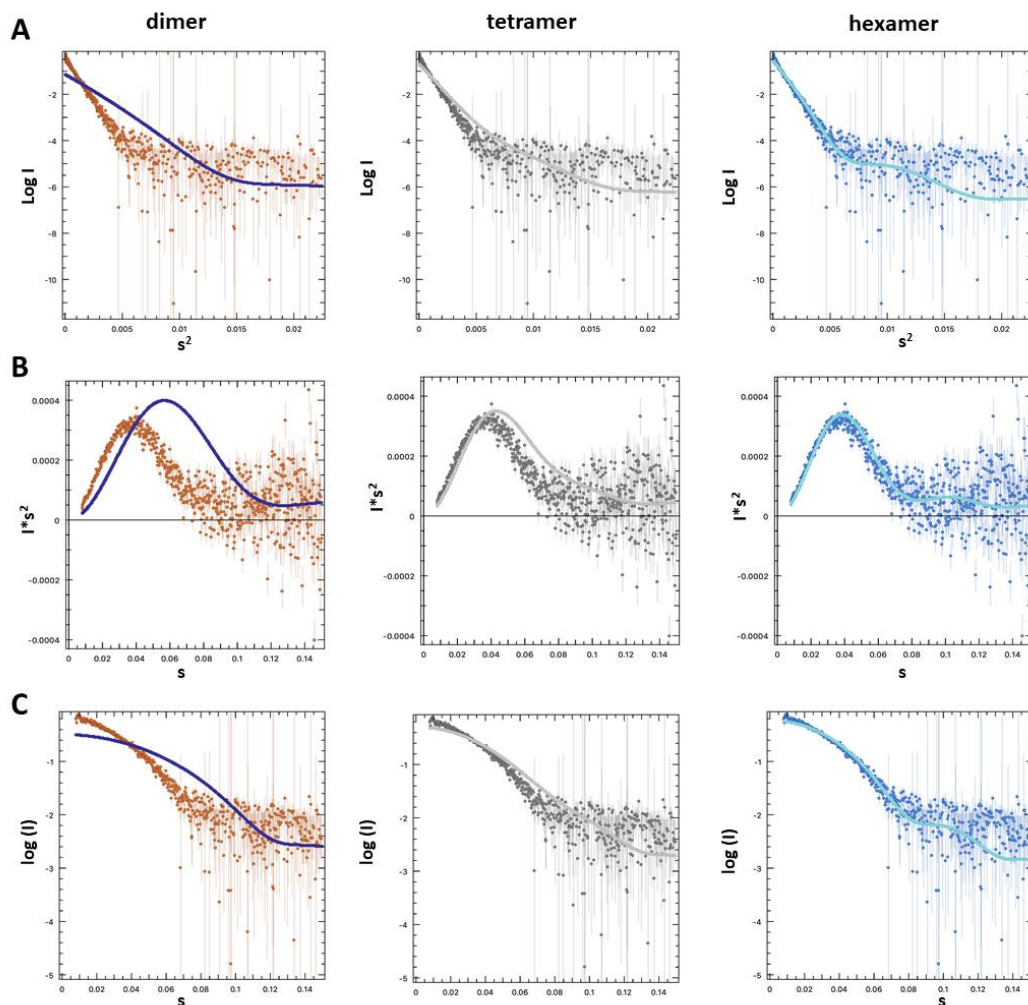
**Figure 7. Overlay of DaHMGR and various class II HMGRs.** Several views containing an overlay of the main body of the protein, excluding the C-terminal domain, of DaHMGR and various other HMGRs. Apo DaHMGR in green (PDB 6eeu), apo PmHMGR in light gray (PDB 4I64), and apo SpHMGR in dark gray (PDB 3QAU), are all shown in cartoon representation. The overall structures align well with no significant structure differences.



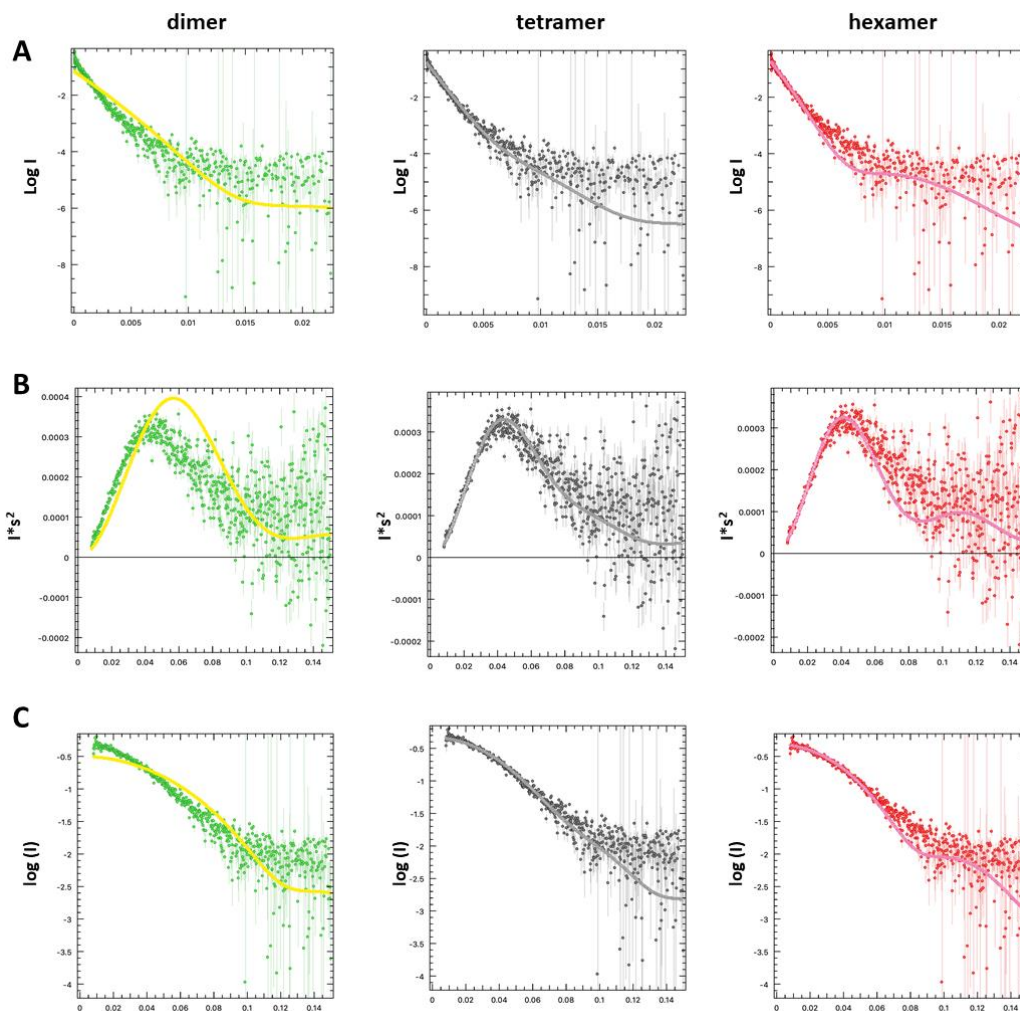
**Figure 8. Size Exclusion Chromatography (SEC) of SpHMGR.** SEC elution profile of SpHMGR. In gray represents protein standards with the peaks corresponding to the size of proteins used. In green represents the elution profile of DaHMGR.



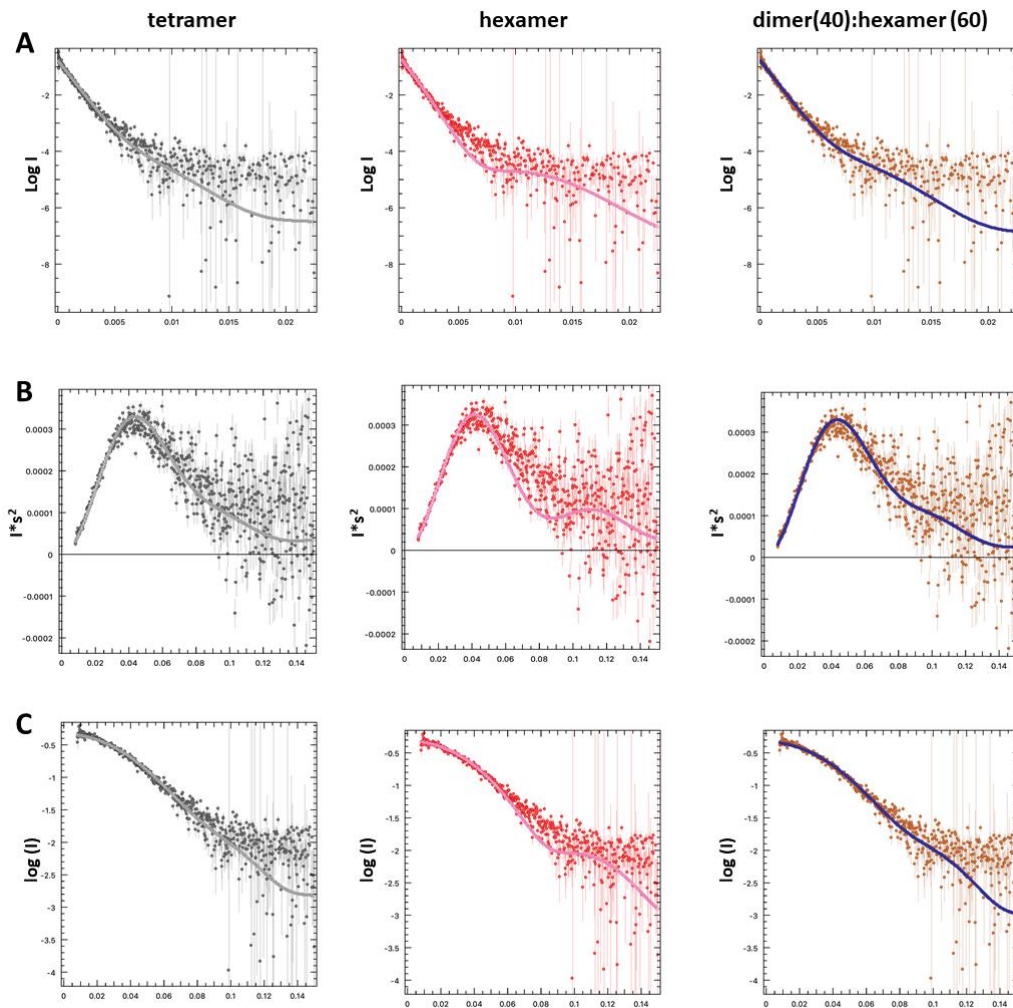
**Figure 9. Size Exclusion Chromatography (SEC) of DaHMGR.** SEC elution profile of DaHMGR. In gray represents protein standards with the peaks corresponding to the size of proteins used. In maroon represents the elution profile of DaHMGR. The y-axis, which represents absorbance at  $A_{280}$  are on the left and right for the protein standards and DaHMGR, respectively.



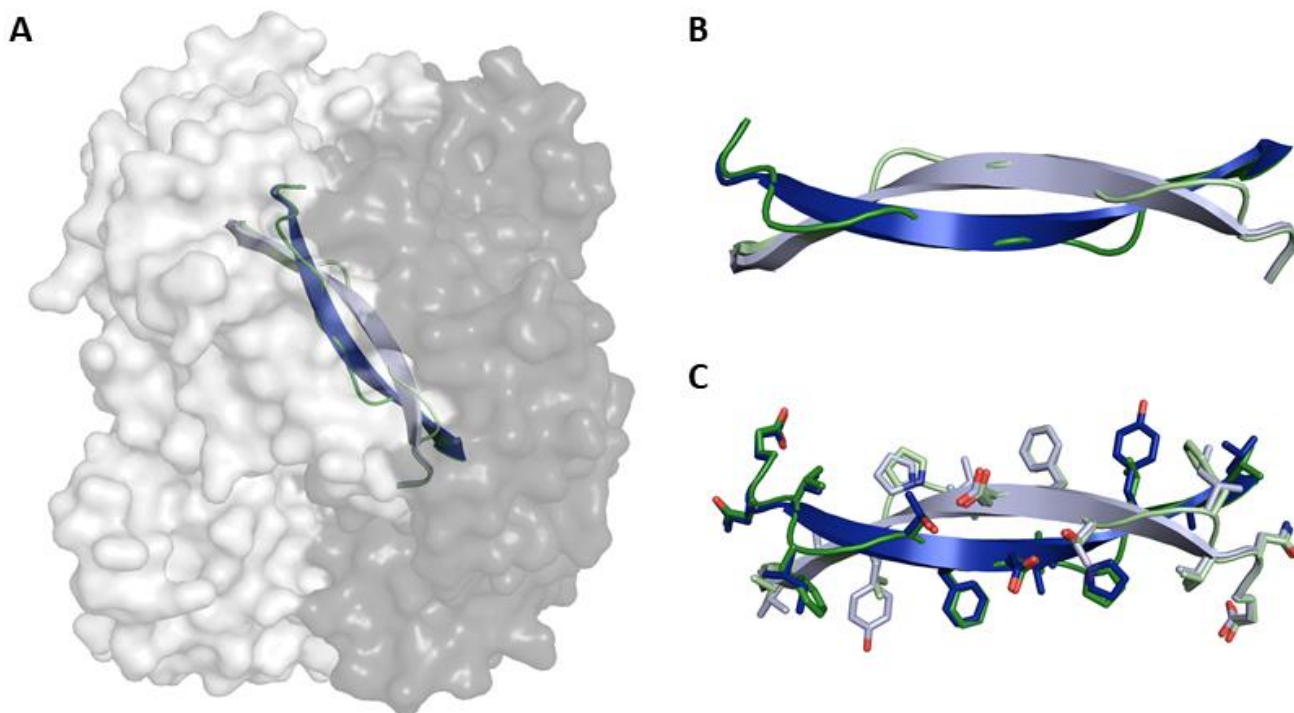
**Figure 10. SAXS analysis of the larger species of DaHMGR.** A) Guinier plot, B) Kratky plot, and C) Logarithmic plot of the larger species eluted in the SEC-SAXS experiment. The dots represent the data, and the fits represent a theoretical fit of the oligomer noted on the top (dimer, tetramer, and hexamer).



**Figure 11. SAXS analysis of the smaller species of DaHMGR. A)** Guinier plot, **B)** Kratky plot, and **C)** Logarithmic plot of the smaller species eluted in the SEC-SAXS experiment. The dots represent the data, and the fits represent a theoretical fit of the oligomer noted on the top (dimer, tetramer, and hexamer).

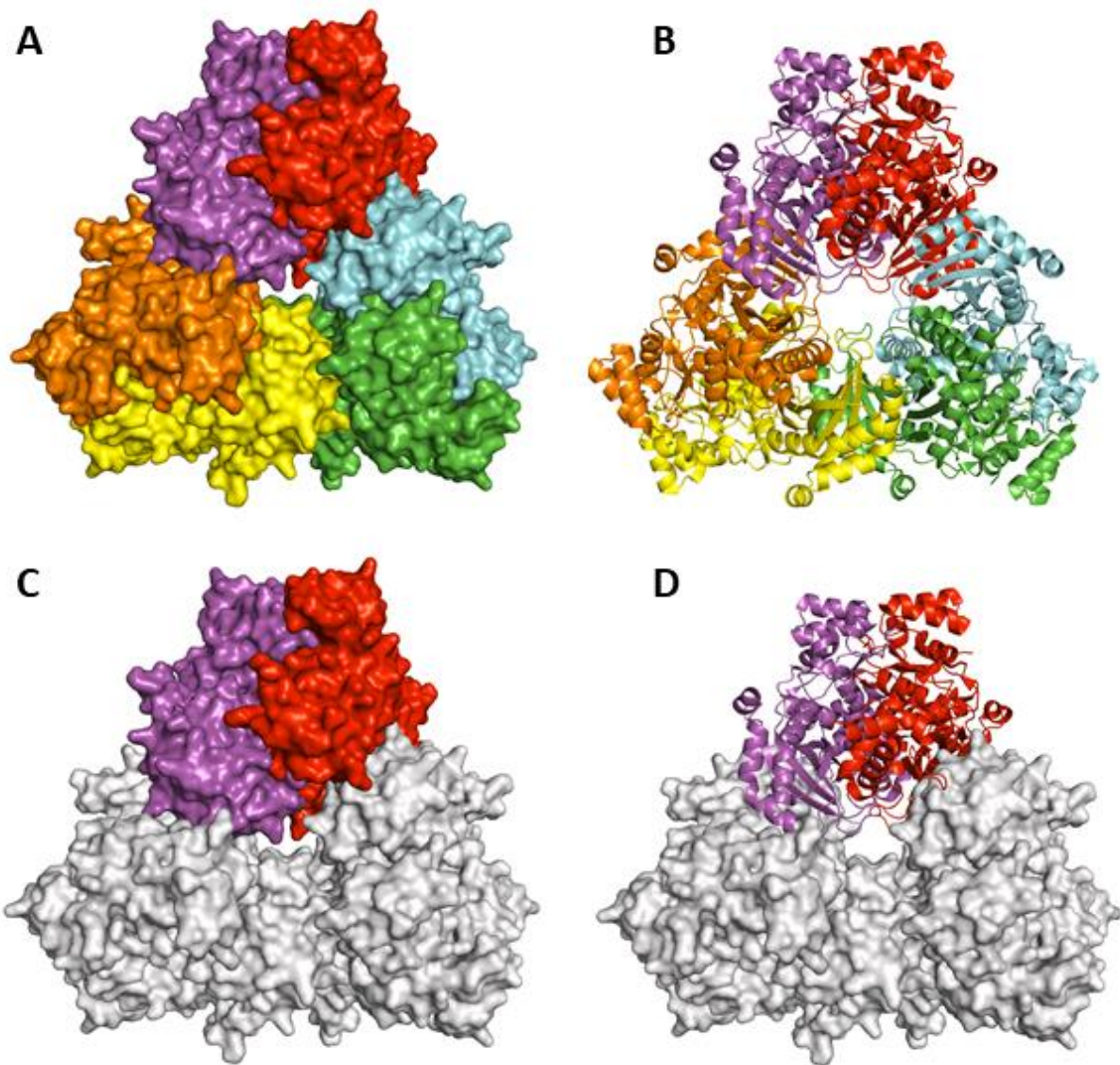


**Figure 12. SAXS analysis of the smaller species of DaHMGR with additional dimer: hexamer fits. A) Guinier plot, B) Kratky plot, and C) Logarithmic plot of the smaller species eluted in the SEC-SAXS experiment. In addition to analyzing the data with respect to a tetramer and hexamer, a dimer: hexamer ratio of 40%:60%, respectively, was used. The dots represent the data, and the fits represent a theoretical fit of the oligomer noted.**

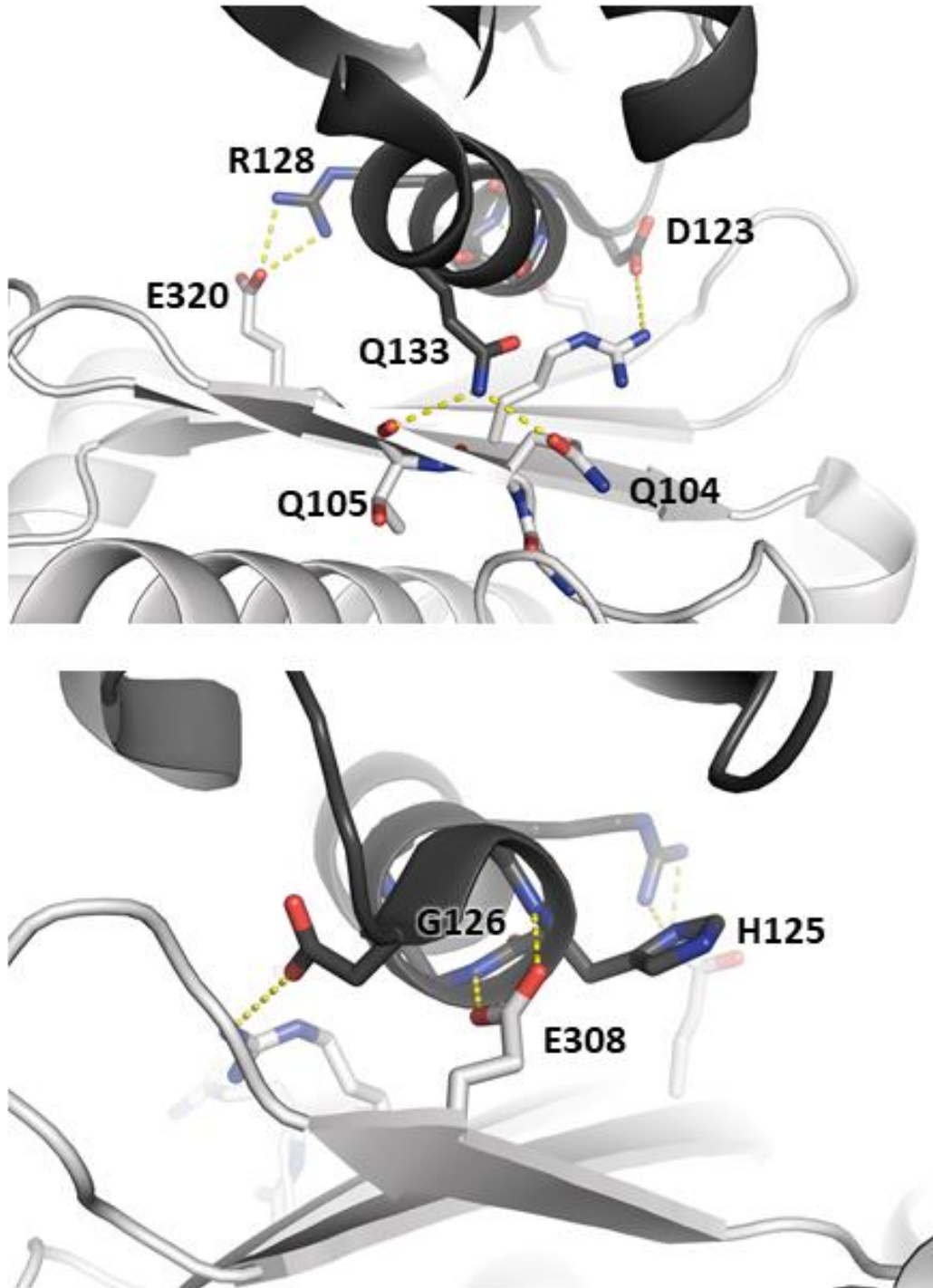


**Figure 13. Comparison between dimerization motif of PmHMGR and DaHMGR.** **A)** Apo PmHMGR (PDB 4I64) as a dimer shown in surface representation with one monomer in light gray and its adjacent monomer in dark gray. The dimerization motif of PmHMGR (dark blue and light blue) and DaHMGR (dark green and light green) are depicted as cartoon. **B)** Close up of an overlay of the dimerization motif shown as a cartoon of PmHMGR and DaHMGR. Dark blue and light blue represent adjacent monomers of PmHMGR, and dark green and light green represent adjacent monomers of DaHMGR. **C)** The sidechains of B are shown as sticks. Coloring of atoms are as follows: O = red, N = blue, C = color of respective cartoon representation. In PmHMGR, the dimerization motif makes up a  $\beta$ -sheet whereas in DaHMGR a strand is observed.

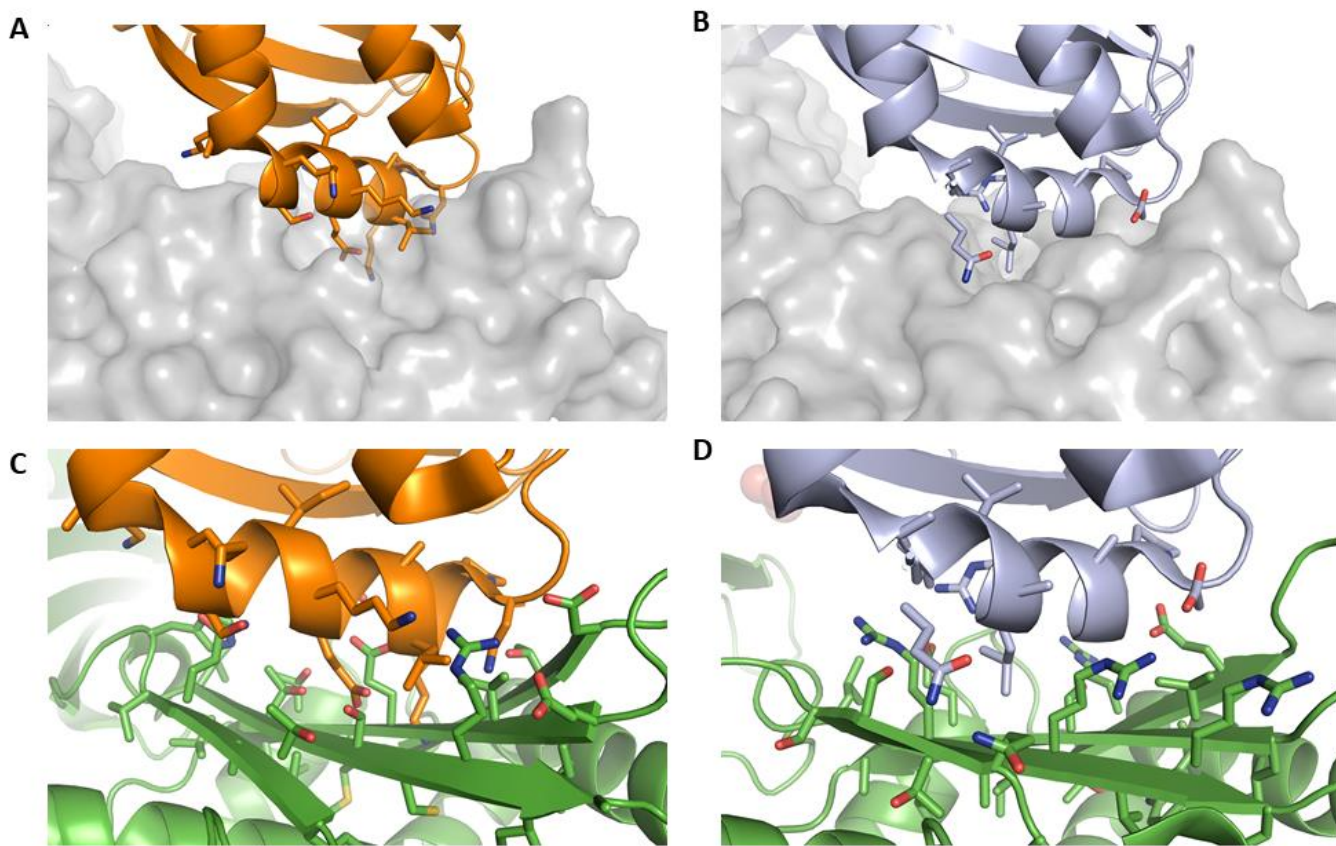




**Figure 14. Hexameric assembly of DaHMGR.** Hexamer, formed through crystallographic symmetry of DaHMGR (PDB 6EEU). Each monomer that makes up the hexamer are colored differently. **A)** Surface representation of hexamer. **B)** Cartoon representation of hexamer. **C)** Surface representation of the hexamer with a dimer colored in purple and red. **D)** The same as C) with the corresponding dimer shown in cartoon representation.



**Figure 15. Hexameric interface of DaHMGR.** Specific interactions between adjacent monomers that constitute the hexameric interface which include interactions proposed by Peacock et al.



**Figure 16. Comparing residues at the hexameric interface in DaHMGR with the aligned sequences of SpHMGR.** **A)** and **C)** represent an overlay of adjacent monomers of SpHMGR with the hexameric interface observed in DaHMGR modeling the nature of this interface if SpHMGR were to form hexamers. **B)** and **D)** represent the hexameric interface of DaHMGR. **A)** and **B)** shown one monomer in cartoon representation with selected residues shown in sticks and the adjacent monomer in surface representation colored gray. **C)** and **D)** show cartoon representation of a both monomers with selected residues shown in sticks. Atom coloring is as follows: O = red, N = blue, C = colored with respect to the cartoon representation. From this, one observes that in SpHMGR (A and C) there exists steric clashes that would not favor the hexamer formation, unlike those observed in DaHMGR (B and D).

taatacgactcactatagggcatATGTCGCCGATTCTCGCCTGCCGAACTTTCGTGCATTGACGCCTGCTCAG  
 CGTCGCGATTTTTTGGCAGATGCGTGTGGATTATCCGATGTCTGAGCGTGCCTGCTGGCTGCCCAGGCGCGCT  
 GCCGCTTGCCCTTGGCGGACGGCATGATCGAAAAATGTGTTGGGAGTTTTGAGCTGCCGTGGGCGTTGCCGGCA  
 ATTTTCGGGTCAATGGACGTGATGTGCTGGTCCCAGTGGCAGTGGAAAGAGCCTTCTGTCGTTGCCGACGCGAGC  
 TATATGGCGAAATTAGCCCGCAGACGGCGGCTTTCAGACATCGTCTACATTGCCACTGATGCGTGCCCAAGT  
 GCAAGTGCTGGGTGCTACTGACCCTCATGGCGCCCGTCTGGCGGTTTTACAAGCGGTGCACAGATCATTGAAC  
 GCGCGAACTCTCGCGATAAAAGTATTAATCGGTTTAGGGGCGGTTGTAAAAGATATCGAGGTACACGTGTTTCCG  
 GATACTCCGCGGGCCCGATGCTGGTGTGCATCTGATTGTTGATGTCCGTGACGCGATGGGTGCCAACACCGT  
 TAATACGATGGCGGAATCGGTGGCCCCGCTGGTTGAAAAAATCACCGTGGCTCTGTCCGTCTGCGCATTCTGA  
 GTAACCTGGCAGATCTGCGCTGGCGCTGCCCGCTGCGTCTGACCCCGCAGACGCTGGCAACCCAAGACCGT  
 AGCGGTGAAGAAATATCGAAGGCGTGTGGATGCGTATACGTTTCGACGCAATCGACCCGTACCGTGTGCAAC  
 CCATAACAAAGGTATTATGAATGGCATCGATCCGGTATTGTTGCCACCGTAATGACTGGCGTGCAGTGGAAAG  
 CAGGCGTACGCGTATGCCAGTGCCTCCGTTTCATACAGTCCCTGACCCGTTGGGAAAAAGATGCAGGTGGC  
 GCTCTGGTTGGTCAATTGAACTGCCGATGCCGTCGGTCTGGTGGGTGGCGCAACCAAAACGCATCCGCTGGC  
 ACGTCTGGCACTGAAAAATCATGGATCTGCAGTCCGCACAGCAACTGGGTGAAATTCGCCGACGCTGTGGGTCTGG  
 CACAGAACCTGGGCGCTCTGCGTGCCTGACTACCGAAGGTATCCAACGCGGCCACATGGCACTGCACGCTCGT  
 AATATTGCGCTGGTTGCGGGTGCCACGGGCGATGAAGTTGACGCGATCGCTCGTCAGCTGGCGGCCGAAACCGA  
 TGTTTCGCACCGACCGTGCCTGGAAGTCTTGGCAGCTCTGCGTGCCTGCGTGGCtaaggatcccgctgagcaata  
actagc

**Codon optimized gene of HMGR from *Delftia acidovorans* (DaHMGR)**

taatacgactcactatagggcatATGAAGATTAGCTGGAATGGTTTTAGCAAGAAGAGCTATCAAGAACGTTCTG  
 GAACTGCTGAAGGCGCAAGCGCTGCTGAGCCCGGAACGCCAGGCGAGCCTGGAGAAGGACGAACAAATGAGCGT  
 GACCGTTGCCGATCAGCTGAGCGAGAACGTGGTTGGCACCTTCAGCCTGCCGTATAGCCTGGTGCCGGAAGTGC  
 TGGTTAACGGTCAAGAATACACCGTGCCTATGTTACCGAGGAACCGAGCGTGGTTGCGGGCGGCGAGCTACGCG  
 AGCAAGATCATTAAACGTGCGGGTGGCTTTACC CGCAGGTTTACCAGCGTCAAATGATCGGTCAGGTGGCGCT  
 GTATCAAGTTGCGAACCCGAAAGCTGGCGCAAGAAAAAATGCGAGCAAGAAAGCGGAGCTGCTGGAACTGGCGA  
 ACCAGGCGTACCCGAGCATCGTGAAGCGTGGTGGCGGTGCGCGTATCTGCACGTTGAGCAAATCAAAGGCGAA  
 CCGGACTTCTGGTGGTTTTACATTCACGTGGATAACCCAGGAAGCGATGGGTGCGAACATGCTGAACACCATGCT  
 GGAAGCGCTGAAGCCGTTCTGGAGGAACTGAGCCAGGGCCAAAGCCTGATGGGTATCCTGAGCAAATATGCGA  
 CCGACAGCCTGGTGACCGGAGCTGCCGTTTTCGTTATCTGAGCCGTCAGAAGGATCAAGGCCGTGAG  
 ATCGCGGAAAAAATGCGCTGGCGAGCCAGTTTGGCGAAGCGGACCCGTATCGTGCAGCGACCCACAACAAAGG  
 CATCTTTAACGGCATCGATGCAATTTCTGATCGCGACGGGCAATGACTGGCGGCCATTGAAGCGGGCGCACACG  
 CCTTTGCGAGCCGCGATGGTCTTATCAAGGTCTGAGCTGCTGGACCCTGGACCTGGAACCGGAGGAACTGGTC  
 GGTGAGATGACGCTGCCGATGCCAGTGGCAACCAAGGTGGTAGCATTGGCCTGAACCCGCGCGTGGCCCTGAG  
 CCATGATCTGCTGGGTAAACCCGTTGCCCCGTTGAACTGGCGCAAATTTATTTGTGAGCATCGGTCTGGCGCAAAAT  
 TCGCGGCACTGAAAGCCCTGGTTAGCACCGGTATTTCAGCAGGGCCACATGAAACTGCAGGCGAAATCTCTGGCA  
 CTGCTGGCCGGCGCATCTGAGAGCGAAGTTGCCCGCTGGTTGAACGTTCTGATTAGCGACAAGACCTTTAACCT  
 GGAGACCGCACAAACGTTACCTGGAATAATCTGCGCAGCtaaggatcccgctgagcaataactagc

**Codon optimized gene of HMGR from *Streptococcus pneumoniae* (SpHMGR)**

**Figure 17. Codon-optimized DaHMGR and SpHMGR genes.** T7 promoter and terminator and restriction sites are shown in addition to the start and stop codons for each gene. Underlined = T7 promoter and terminator sequences. **NdeI** and **BamHI** are restriction sites, **ATG** and **TAA** are start and stop codons respectively.

MGSSHHHHHDYDIPTTENLYFQGHMVADSRLPNFRALTPAQRDFLADACGLSDAERALLAAPGALPLALADG  
MIENVFGSFELPLGVAGNFRVNGRDVLPMAVEEPSVAAAASYMAKLAREDDGGFQTSSTLPLMRAQVQVLGVTD  
PHGARLAVLQARAQI IERANSRDKVLI GLGGGCKDIEVHVFPDTPRGPMLVVHLIVDVRDAMGANTVNTMAESV  
APLVEKITGGSVRLRILSNLADLRLARARVRLTPQTLATQDRSGEEI IEGVLDAYTFAAIDPYRAATHNKGIMN  
GIDPVIIVATGNDWRAVEAGAHAYASRSGSYTSLTRWEKDAGGALVGSIELPMPVGLVGGATKTHPLARLALKIM  
DLQSAQQQLGEIAAAVGLAQN LGALRALATEGIQRGHMALHARNIALVAGATGDEVDAVARQLAAEHDVRTDRAL  
EVLAAALRARA

Amino acid sequence of HMGR from *Delftia acidovorans* (DaHMGR)

MW: 48.1 kDa

Ext. Coeff: 21,500 M<sup>-1</sup> cm<sup>-1</sup>

pH: 6.24

MGSSHHHHHDYDIPTTENLYFQGHMKISWNGFSKKS YQERLELLKAQALLSPERQASLEKDEQMSVTVADQLS  
ENVVGTFSLPYSLVPEVLVNGQEYTPYVTEEPSVAAAASYASKI IKRAGGFQAQVHQRMIGQVALYQVANPK  
LAQEKIASKKAELLELANQAYPSIVKRGGGARDLHVEQIKGEPDFLVVYIHVDTQEAMGANMLNTMLEALKPVL  
EELSQQSIMGILSNYATDSLVTASCRIFRYLSRQKQDGREIAEKIALASQFAQADPYRAATHNKGIFNGIDA  
IL IATGNDWRAIEAGAHAFASRDGRYQGLSCWTLDLEREELVGEMTLPMPVATKGGSIGLNPRVALSHDLLGNP  
SARELAQIIVSIGLAQNFAALKALVSTGIQQGHMKLQAKSLALLAGASESEVAPLVERLISDKTFNLETAQRYL  
ENLRS

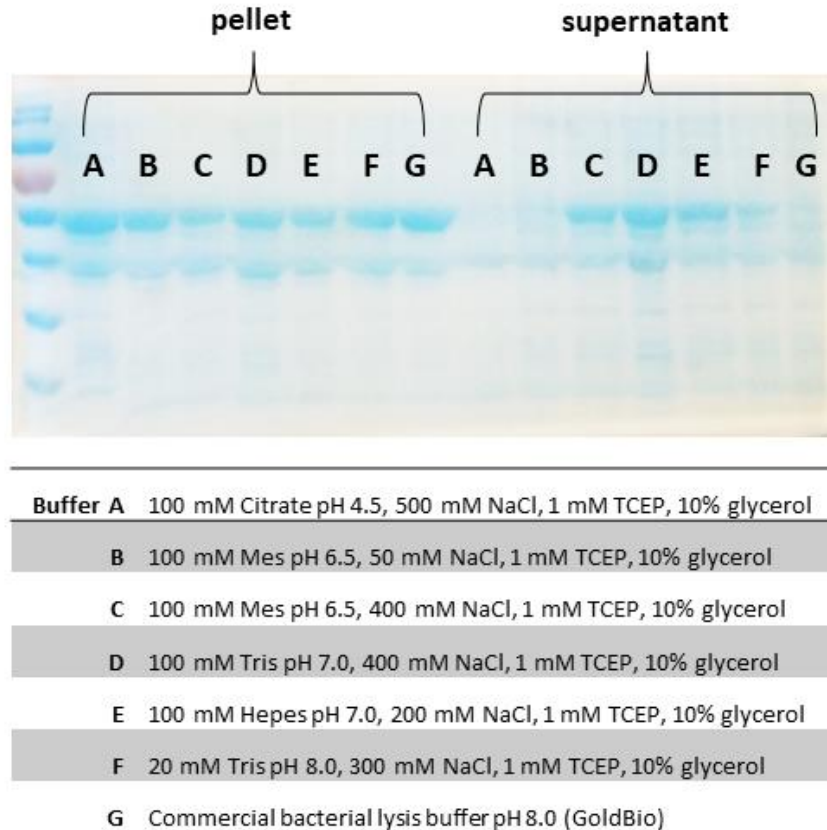
Amino acid sequence of HMGR from *Streptococcus pneumoniae* (SpHMGR)

MW: 49.2 kDa

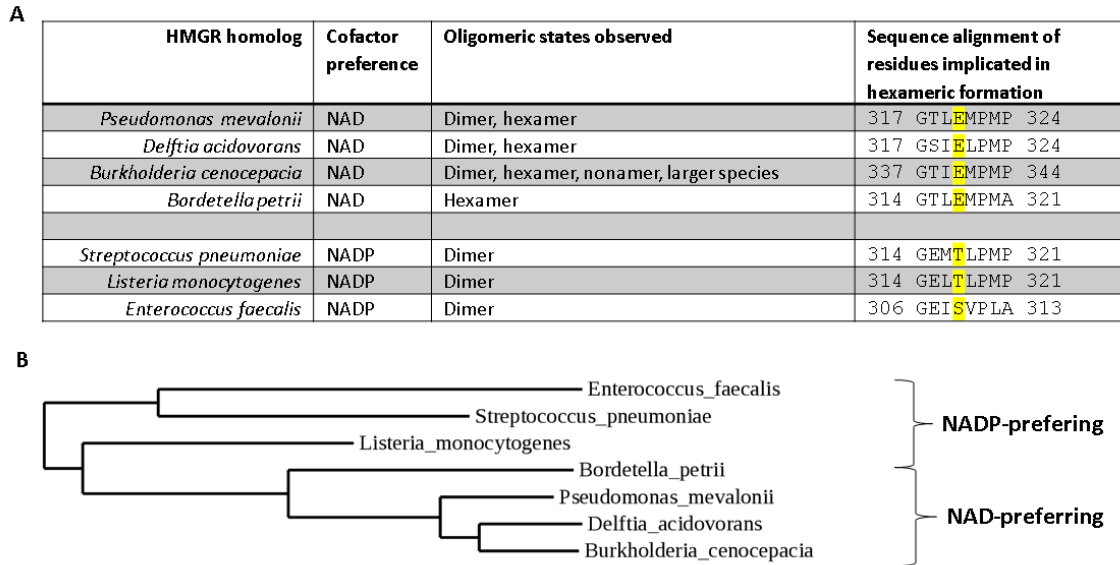
Ext. Coeff: 38,900 M<sup>-1</sup> cm<sup>-1</sup>

pH: 6.08

**Figure 18.** DaHMGR and SpHMGR amino acid sequences. Underlined is the TEV protease cleavage site which proceeds the hexa-histidine tag used for purification during Ni-NTA affinity chromatography. Underlined = TEV protease cleavage site.



**Figure 19. Lysis buffer screening conditions for DaHMGR purification.** On the bottom are conditions A through G which vary by type of buffering reagent, pH, and salt content. On top is an SDS-PAGE gel of the pellet versus supernatant fractions when cells containing overexpressed DaHMGR was lysed and centrifuged using the respective buffers.



**Figure 20. Oligomerization observed in various characterized HMGRs.** **A)** A table showing the relationship between cofactor preference and oligomerization among class II HMGRs. Additionally, a sequence alignment of a region encompassing a glutamate residue that is conserved among NAD(H)-preferring HMGRs which form hexameric interface. This glutamate interacts with an arginine for the adjacent monomer and is involved in stabilizing hexameric interface. In NADP(H)-preferring HMGRs, this glutamate is replaced with smaller residues. **B)** A phylogenetic tree created with PhyML showing how cofactor preference is manifested in various HMGR homologs.

**Table 1. X-ray Data Collection and Refinement Statistics.**

PDB ID	apo DaHMGR 6EEU
<b>Data Collection</b>	
Beamline	APS, 24-ID-E
Wavelength (Å)	0.9792
Space group	<i>P</i> 321
Unit cell	
<i>a=b, c</i> (Å)	100.54, 75.86
Resolution (Å)	57.20–1.93 (1.99–1.93)
Wilson B (Å <sup>2</sup> )	20.59
Total reflections	138,912 (9,597)
Unique reflections	33,218 (2,243)
Multiplicity	4.2 (4.3)
Completeness (%)	98.9 (100.0)
Mean <i>I</i> / $\sigma$ ( <i>I</i> )	8.1 (2.0)
<i>R</i> merge	0.117 (0.823)
<i>R</i> meas	0.134 (0.941)
CC1/2	0.994 (0.602)
<b>Refinement</b>	
<i>R</i> work	0.1597 (0.2325)
<i>R</i> free	0.1980 (0.2845)
r.m.s.d. bonds (Å)	0.006
r.m.s.d. angles (°)	0.794
Number of atoms	3,097
Average B factors (Å <sup>2</sup> )	
All atoms	22.95
Protein	21.7
NADH/citrate/mevalonate	N/A
Water	31.76
Ramachandran analysis	
Favored (%)	98.12
Allowed (%)	1.88
Outliers (%)	0
MolProbity Clashscore	1.62

---

Statistics for the highest-resolution shell shown in parentheses.



## 2.8. References

- [1] A. W. Senior *et al.*, “Improved protein structure prediction using potentials from deep learning,” *Nature*, vol. 577, no. 7792, pp. 706–710, Jan. 2020, doi: 10.1038/s41586-019-1923-7.
- [2] J. Jumper *et al.*, “Highly accurate protein structure prediction with AlphaFold,” *Nature*, Jul. 2021, doi: 10.1038/s41586-021-03819-2.
- [3] R. W. Newberry and R. T. Raines, “Secondary Forces in Protein Folding,” *ACS Chem. Biol.*, vol. 14, no. 8, pp. 1677–1686, Aug. 2019, doi: 10.1021/acscchembio.9b00339.
- [4] M. Q. Fatmi and C. A. Chang, “The Role of Oligomerization and Cooperative Regulation in Protein Function: The Case of Tryptophan Synthase,” *PLoS Comput Biol*, vol. 6, no. 11, p. e1000994, Nov. 2010, doi: 10.1371/journal.pcbi.1000994.
- [5] C. Lawrence, V. Rodwell, and C. Stauffacher, “Crystal structure of *Pseudomonas mevalonii* HMG-CoA reductase at 3.0 angstrom resolution,” *Science*, vol. 268, no. 5218, pp. 1758–1762, Jun. 1995, doi: 10.1126/science.7792601.
- [6] L. Taberner, D. A. Bochar, V. W. Rodwell, and C. V. Stauffacher, “Substrate-induced closure of the flap domain in the ternary complex structures provides insights into the mechanism of catalysis by 3-hydroxy-3-methylglutaryl-CoA reductase,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 13, pp. 7167–7171, Jun. 1999, doi: 10.1073/pnas.96.13.7167.
- [7] B. R. Miller and Y. Kung, “Structural Features and Domain Movements Controlling Substrate Binding and Cofactor Specificity in Class II HMG-CoA Reductase,” *Biochemistry*, vol. 57, no. 5, pp. 654–662, Feb. 2018, doi: 10.1021/acs.biochem.7b00999.
- [8] E. S. Istvan, M. Palnitkar, S. K. Buchanan, and J. Deisenhofer, “Crystal structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis,” *EMBO J*, vol. 19, no. 5, pp. 819–830, Mar. 2000, doi: 10.1093/emboj/19.5.819.
- [9] B. E. Haines, O. Wiest, and C. V. Stauffacher, “The Increasingly Complex Mechanism of HMG-CoA Reductase,” *Acc. Chem. Res.*, vol. 46, no. 11, pp. 2416–2426, Nov. 2013, doi: 10.1021/ar3003267.
- [10] R. A. Laskowski, E. G. Hutchinson, A. D. Michie, A. C. Wallace, M. L. Jones, and J. M. Thornton, “PDBsum: a Web-based database of summaries and analyses of all PDB structures,” *Trends Biochem Sci*, vol. 22, no. 12, pp. 488–490, Dec. 1997, doi: 10.1016/s0968-0004(97)01140-7.
- [11] S. Jones and J. M. Thornton, “Protein-protein interactions: A review of protein dimer structures,” *Progress in Biophysics and Molecular Biology*, vol. 63, no. 1, pp. 31–65, 1995, doi: 10.1016/0079-6107(94)00008-W.

- [12] K. Hashimoto and A. R. Panchenko, “Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 47, pp. 20352–20357, Nov. 2010, doi: 10.1073/pnas.1012999107.
- [13] E. van Dijk, A. Hoogeveen, and S. Abeln, “The Hydrophobic Temperature Dependence of Amino Acids Directly Calculated from Protein Structures,” *PLoS Comput Biol*, vol. 11, no. 5, p. e1004277, May 2015, doi: 10.1371/journal.pcbi.1004277.
- [14] M. S. Brown, S. E. Dana, J. M. Dietschy, and M. D. Siperstein, “3-Hydroxy-3-methylglutaryl Coenzyme A Reductase,” *Journal of Biological Chemistry*, vol. 248, no. 13, pp. 4731–4738, Jul. 1973, doi: 10.1016/S0021-9258(19)43725-3.
- [15] R. B. Peacock *et al.*, “Structural and Functional Characterization of Dynamic Oligomerization in *Burkholderia cenocepacia* HMG-CoA Reductase,” *Biochemistry*, vol. 58, no. 38, pp. 3960–3970, Sep. 2019, doi: 10.1021/acs.biochem.9b00494.
- [16] S. M. Ma *et al.*, “Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases,” *Metabolic Engineering*, vol. 13, no. 5, pp. 588–597, Sep. 2011, doi: 10.1016/j.ymben.2011.07.001.
- [17] E. K. Jaffe, “Morpheesins – a new structural paradigm for allosteric regulation,” *Trends in Biochemical Sciences*, vol. 30, no. 9, pp. 490–497, Sep. 2005, doi: 10.1016/j.tibs.2005.07.003.
- [18] A. E. Theivagt, E. N. Amanti, N. J. Beresford, L. Taberero, and J. A. Friesen, “Characterization of an HMG-CoA Reductase from *Listeria monocytogenes* That Exhibits Dual Coenzyme Specificity,” *Biochemistry*, vol. 45, no. 48, pp. 14397–14406, Dec. 2006, doi: 10.1021/bi0614636.
- [19] J. Pérez and Y. Nishino, “Advances in X-ray scattering: from solution SAXS to achievements with coherent beams,” *Current Opinion in Structural Biology*, vol. 22, no. 5, pp. 670–678, Oct. 2012, doi: 10.1016/j.sbi.2012.07.014.
- [20] D. A. Jacques and J. Trewhella, “Small-angle scattering for structural biology-Expanding the frontier while avoiding the pitfalls: Small-Angle Scattering for Structural Biology,” *Protein Science*, vol. 19, no. 4, pp. 642–657, Apr. 2010, doi: 10.1002/pro.351.
- [21] J. L. Cole, J. W. Lary, T. P. Moody, and T. M. Laue, “Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium,” in *Methods in Cell Biology*, vol. 84, Elsevier, 2008, pp. 143–179. doi: 10.1016/S0091-679X(07)84006-4.
- [22] A. Dereeper *et al.*, “Phylogeny.fr: robust phylogenetic analysis for the non-specialist,” *Nucleic Acids Research*, vol. 36, no. Web Server, pp. W465–W469, May 2008, doi: 10.1093/nar/gkn180.

## Chapter III: Kinetic, and structural characterization of HMGR from *Delftia acidovorans* and *Streptococcus pneumoniae*

### 3.1. Summary

After exploring the native state and global structural dynamics of HMGR from *Delftia acidovorans* (DaHMGR) and *Streptococcus pneumoniae* (SpHMGR) in Chapter Two, we now venture into greater depth to further probe the function of the enzyme on an atomic level by exploring the finer details of catalysis. This chapter will begin to unravel the specifics of the mechanism as revealed by structural and functional characteristics of HMGR through studying the X-ray crystallographic crystal structures bound with various ligands as well as the steady-state kinetic parameters of substrate and cofactor. There will be specific focus geared towards the role of the C-terminal flap domain (CTD), introduced in Chapter Two, and how it facilitates HMGR catalysis. Attention will be given to the numerous positions the CTD assumes in various apo and ligand-bound states to substantiate and expand our understanding of the role that this domain has in the mechanism.

This research yielded the first crystal structures of DaHMGR, in apo form as well as ligand-bound forms, complemented by steady-state kinetics revealing this enzyme as an NADH-preferring HMGR and depicting a novel “flipped” conformation of the CTD. Novel structures of DaHMGR and SpHMGR, in addition to mutant constructs of SpHMGR, that are discussed in Chapter Four, will be partially incorporated into this chapter only as it pertains to the role of the CTD since they provide additional insight into this domain. The relatively high-resolution crystal structures will confirm the importance of conserved cofactor and substrate binding site

architecture. They will also corroborate prior research detailing unique observations found in NADH-preferring HMGRs and will reveal a novel location of the C-terminal flap domain which has direct implications for catalysis. With relatively high-resolution structures, greater detail into HMG-CoA catalysis will inform both the order of reactivity as well as the roles of various residues involved. In the end, this chapter will propose a modified reaction mechanism as well as validate the importance of structural motifs that dictate cofactor specificity.

Portions of this chapter were published in:

E. R. Ragwan, E. Arai, and Y. Kung, "New Crystallographic Snapshots of Large Domain Movements in Bacterial 3-Hydroxy-3-methylglutaryl Coenzyme A Reductase," *Biochemistry*, vol. 57, no. 39, pp. 5715–5725, Oct. 2018, doi: 10.1021/acs.biochem.8b00869.

### **3.2. Introduction**

HMGR is classified as an oxidoreductase, more specifically, an NAD(P)-dependent oxidoreductase. Oxidoreductases belong to a large class of enzymes that encompasses almost one third of all enzymatic activities and utilizes a range of both organic and inorganic substrates such as alcohols, amines, and metals [1]. These enzymes perform redox reactions as follows:  $A^- + B \rightarrow A + B^-$  where A represents a substrate that is oxidized during the reaction while B represents a substrate that is reduced; depending on the enzyme, the reverse reaction can also be catalyzed [1]. NAD(P)-dependent oxidoreductases, can reduce or oxidize a substrate by facilitating the donating or accepting of a hydride, respectively, using the cofactor, NAD(P)(H), as the hydride reservoir.

Oxidoreductases typically have a catalytic domain and a cofactor binding domain. The catalytic domain, which binds the substrate, assumes distinctive active site architecture depending

on the enzyme's substrate. The cofactor binding domain, however, often comprises common structural motifs, such as the Rossmann fold, which allow these enzymes to accommodate the nucleotide binding domain of NAD(P). In Chapter Four we will discuss the specifics of cofactor binding as it pertains to the structural nuances that govern the between NADH and NADPH, which in the case of HMGR involves a non-Rossmann fold. However, in this section, some attention will be given to the nature of cofactor binding in the context of the mechanism of catalysis.

In 2013, Haines et al. published "*The Increasingly Complex Mechanism of HMG-CoA Reductase*" and emphasized that despite the diversity of tools utilized to study this enzyme, ranging from biochemical to computational techniques, the details and certainty of the mechanism remains incomplete [2]. This complexity is due to the nature of the reaction where HMGR performs a four-electron reduction of HMG-CoA to mevalonate, which requires two cofactors, multiple intermediates, and a means to control the steps of the sequence of the reaction. The overall mechanism is as follows: HMG-CoA binds and is reduced to mevaldyl-CoA using one equivalent of NAD(P)H. Then, mevaldyl-CoA is cleaved to form mevaldehyde and CoA, while NAD(P)<sup>+</sup> and CoA are removed, as a second NAD(P)H binds to facilitate conversion of mevaldehyde to mevalonate, after which NAD<sup>+</sup> and the final product, mevalonate, can be released.

Since 2013, further insight into the mechanism has been expounded on, and will be discussed here, but there are still questions that persist about the order of events and contributions of residues and conformations. In this section, we will highlight 1) the active site architecture and mechanistic proposals, and 2) the C-terminal flap domain (CTD) in its role during the course of the reaction. In doing so, we will illustrate how the research presented here provides marked improvements in our understanding of this enzyme, but we will also gain an appreciation for the uncertainties that remain. Since research into HMGR falls at the juncture between evaluating

existing work and the need for further investigation, HMGR continues to be a relevant case study in enzymology, especially in the field of oxidoreductases, that promises fruitful inquiry.

### 3.2.1. Active site architecture and mechanism

The sequence identities within class II HMGRs are ~50%, and contain conserved active site residues, and as pointed out in Chapter Two, possess similar overall folds [2]. Figure 1 shows the sequence alignment of a variety of class II HMGRs, noting the highly conserved regions. The active site, made up of two monomers, is at the apex of a “V” shape where the substrate of one monomer and cofactor of the adjacent monomer meet at a buried core that exists at the homodimeric interface and extend towards the solvent in different directions, as discussed in Chapter Two [3]. The conserved active site residues have been the center of attention with regards to proposed mechanisms. In early studies, done in yeast HMGR, unspecified acidic and basic residues were suggested as being crucial for catalysis [2], [4]. Additional work highlighted the importance of a yet-to-be identified acidic residue and a cationic histidine [2], [5]. Based on the sequence alignment of class II HMGRs and with the knowledge that a conserved acidic and histidine residue was important for catalysis, site-directed mutagenesis studies homed in on Glu83 and His381 [2], [6], [7].

The advancement in biochemical tools contributed to acquiring greater details about the mechanism. The first crystal structure, described in Chapter Two, showed a well-defined HMG pocket to accommodate the substrate and revealed a “shallow” surface groove with fewer stabilizing interactions with the CoA moiety [2]. Additional higher resolution structures of PmHMGR describe specific hydrogen bonds and water mediated interactions involved in substrate binding. For example, in a structure of HMG-CoA-bound PmHMGR, Steussy et al. highlights

hydrogen bonds between the protein and both ends of HMG-CoA [8]. In their structure, they propose that at one end of the HMG-CoA binding pocket Arg11 makes a hydrophobic interaction between the CoA adenine and its sidechain carbons and they observe a hydrogen bond between the imine side group and the adenine phosphate [8]. At the well-defined HMG pocket, on the other end, two hydrogen bonds between Arg261 imine groups interact with the carboxyl group of HMG, and between both ends, the authors suggest that there are water-mediated interactions involved, albeit affording torsional flexibility in the central region [8].

Similar types of interactions that anchor bond ends of HMG-CoA are observed in HMG-CoA-bound SpHMGR where Ser9 and Gly7 make backbone and side chain interactions between the 3'-phosphate of the adenosine ribose, and Arg257 coordinates the carboxylate at the other end [3]. In addition, Ser85 of PmHMGR has been shown to stabilize reaction intermediates by interacting with the first amide bond of CoA [8]. Figure 2 depicts both ends of the binding pocket of HMG-CoA highlighting some of the important direct interactions between HMG-CoA and residues of PmHMGR and SpHMGR.

There are also conserved regions within the cofactor binding site that pertain to cofactor specificity. These interactions will be discussed in detail in Chapter Four but for the purpose of the reaction mechanism there are additional noteworthy residues that are implicated in catalysis. At the active site, well conserved asparagine, lysine and aspartate residues typically serve to coordinate the region around the nicotinamide ring of the cofactor and have been shown to be important for catalysis [2], [3]. More specifically, as observed in NADPH-bound SpHMGR, Asn212 hydrogen bonds with the amide group of the nicotinamide ring, and Asp279 hydrogen bonds with the 2-hydroxyl group of the nicotinamide's ribose [3]. As pointed out in the HMG-

CoA-bound PmHMGR, Lys267 is positioned to be able to stabilize the negative charge intermediates including carboxylate groups from the active site Glu and Asp mentioned above [2]. In total, there have been several proposed mechanisms that have developed chronologically based on the incremental knowledge gained from additional studies.

### 3.2.2. Role of the C-terminal domain in catalysis

As pointed out in the overview of the reaction, the series of steps involved in the reaction is complex. Substrate and cofactor must bind to generate the first intermediate, mevaldyl-CoA, followed by the exchange of the oxidized cofactor with a reduced one without the release of the mevaldyl-CoA. Once mevaldyl-CoA is converted to mevaldehyde, HS-CoA needs to be released while mevaldehyde needs to be retained. Finally, mevaldehyde is converted to the product, mevalonate, and both the product and the second used cofactor are released. It is not known, for example, when HS-CoA is released, and the precise order remains elusive.

Of importance is the fact that HMGR intermediates mevaldyl-CoA and mevaldehyde are not released during the catalytic cycle [9], suggesting a mechanism by which these intermediates are retained in the active site until mevalonate is formed. In contrast, the cofactors NAD(P)H and NAD(P)<sup>+</sup> do need to enter and exit at specific times within the reaction cycle, suggesting a mechanism that timely controls cofactor exchange. Taken together, these two important mechanistic features of retaining HMG-CoA intermediates while modulating the entry and exit of cofactors provide the rationale for what is observed in HMGR crystal structures, where a flexible C-terminal domain (CTD) that is near both the substrate and cofactor sites acts as a flap to help, at least in part, to control these steps. This flexible CTD flap completes a fascinating feat of engineering encompassing a range of motion that gives it the ability to cover either or both of the



substrate and cofactor sites in a specific order that regulates the entry and exist of the respective ligand.

The initial insights regarding the CTD was gained through the work done by Taberno et al. where they determined the structures of two non-productive ternary complexes, PmHMGR with HMG-CoA/NAD<sup>+</sup> and mevalonate/NADH of which the former structure depicted the CTD covering the active site [10]. The authors described the former structure as being involved in a substrate-induced closing of the CTD. In this closed conformation, the CTD ensures that the catalytic His381, which was previously identified as a catalytically important cationic residue, is positioned appropriately in the active site. In addition, the authors propose that as a result of the closed conformation, a hydrophobic pocket, made up of residues Ile213, Leu371, and Ile377, covers the hydride transfer reaction, thereby sheltering it from the solvent [10]. The closed conformation is further stabilized by interactions between residues within the CTD of one monomer and the small domain of the adjacent monomer[10]. Figure 3 shows the position of the CTD as it covers the active site, along with the accompanying positioning of the catalytically important histidine that is enabled by the CTD adopting this closed conformation.

A structure of apo SpHMGR (PDB 3QAU), without an associated publication, was deposited in the Protein Database, and upon inspection of the structure revealed a novel location of the CTD. In contrast to the closed conformation found in the above ternary complex of PmHMGR with HMG-CoA/NAD<sup>+</sup> that covers the active site, this apo structure revealed an open conformation where the CTD was flipped away from the active site[3]. This conformation repositions the catalytically important histidine away from the active site. In a subsequent structure determined by the Kung Lab of NADPH-bound SpHMGR, two additional open conformations

were observed[3]. Neither of the open conformations align directly over each other and it was speculated that crystal contacts with adjacent molecules may have given rise to these slightly varied positions [3]. However, the take-away from these various open and closed conformations is that it appears that the CTD flap domain closes for the purpose of positioning the histidine into the active site, sheltering the hydride transfer reaction and when opened cofactor exchange can occur. Figure 4 shows the various open conformations observed till this point.

Yet another location of the CTD was observed in an HMG-CoA-bound SpHMGR that was described as being in a partially closed conformation [3]. Whereas in the closed conformation the CTD covered the active site, and in the open conformation it was flipped away, this partially closed conformation located the CTD away from the cofactor-binding site but still covering the substrate-binding site [3]. In the closed conformation, with both substrate and cofactor bound in a dead-end complex, the CTD covered the entire active site, including both substrate and cofactor sites. This contrasts with the SpHMGR structure that has only substrate bound which reveals the partially closed conformation that only covers the substrate site.

This partially closed conformation would allow cofactor exchange without the release of substrate intermediates, which corroborates prior studies that show that reaction intermediates are not released during the course of the reaction. Figure 5 shows all the conformations of the CTD determined from PmHMGR and SpHMGR structures. What is revealed through all these structures showing the CTD in various location is that the CTD is mobile, spanning various locations for the integrity of the reaction. While the order of events remains unclear, an evolving paradigm of how the CTD facilitates the reaction, that is optimized by preventing intermediates from being released, and the mechanism by which two equivalents of cofactors are used in this sequence, is becoming

clearer, but further work will continue to expand our understanding of this fascinating, dynamic enzyme.

### 3.2.3. Goals of this section

As observed in the above chronological unraveling of the mechanism of HMGR, additional studies done on different homologs incrementally added and expounded on prior assumptions. Thus, this work serves to add to the growing body of knowledge that can both corroborate prior research and extend our understanding. For this reason, we utilize a previously less characterized HMGR, from *Delftia acidovorans* (DaHMGR), with the desire to provide novel crystal structures and complementary biochemical data to further our understanding of class II HMGR. The vision was birthed out of the need to clarify the importance of active site residues and to understand the role of the CTD more fully as it traverses through different conformations during the course of the reaction.

More specifically, one goal sought to characterize DaHMGR using steady-state kinetics and X-ray crystallography. Several HMGRs have been kinetically characterized as revealed in Table 4 which outlines a variety of kinetic parameters determined for HMGRs that have been published thus far. This table provides the following: the organism from which the HMGR comes from; and any steady-state kinetic parameters obtained for substrate, product, or cofactor. However, prior studies on DaHMGR are limited in biochemical characterization. Of the limited published biochemical data, DaHMGR was featured in one study using engineered *E. coli* strains where HMGRs from several different species were expressed alongside the remaining MEV pathway enzymes from *Saccharomyces cerevisiae*. The pathway was engineered to produce amorphaadiene, the isoprenoid precursor to the antimalarial drug artemisinin, and out of the HMGR

homologs tested, the greatest yield was achieved when using DaHMGR [11]. Despite this intriguing observation, a more thorough kinetic characterization of the steady-state kinetics using HMG-CoA and NAD(P)H, is lacking. Thus, this goal seeks to utilize steady-state activity assays to assess substrate and cofactor affinity and in doing so hope to determine its cofactor preference.

With the knowledge of its cofactor preference, the next step in this goal of characterizing DaHMGR seeks to crystallize DaHMGR with its preferred cofactor to observe the interactions with the cofactor. Prior to this work, structures of only two class II HMGRs, PmHMGR and SpHMGR, were determined. These two structures, one of the NADH-preferring PmHMGR and one of the NADPH-preferring SpHMGR, have provided some structural basis for cofactor specificity. By determining the structure of DaHMGR with its preferred cofactor, one is able to assess prior hypotheses about the structural determinants that confer cofactor specificity. This will aid the work done in Chapter Four that aims to utilize this knowledge to generate modified HMGRs with modified cofactor preferences.

A second goal of this section involved solving crystal structures of DaHMGR and SpHMGR that capture the conformation of the CTD. By locating the CTD, with or without ligands bound, this study will add to the limited number of structures that depict the CTD in their models. The hope is that we will eventually be equipped with the understanding of how the CTD behaves as it accommodates and facilitates each step of the mechanism. Incidentally, while studying cofactor specificity, which is the focus of Chapter Four, additional SpHMGR mutant constructs generated crystal structures containing additional insight related to the C-terminal domain (CTD). Since the CTD is one of the focuses of the present chapter, the crystal structures from these SpHMGR mutant constructs will be partly incorporated into the Discussion of this chapter but only as they pertain to the C-terminal domain (CTD), with the hope of expanding our

understanding of the role of the CTD during the mechanism. In sum, this section continues from Chapter Two in characterizing HMGR but specifically exploring the mechanism utilizing structural and functional data.

### 3.3. Results

#### 3.3.1. Activity assays

The activity of DaHMGR was monitored by the decrease in absorbance at  $A_{340}$  nm since the oxidation of NAD(P)H to NAD(P)<sup>+</sup> causes a loss of absorbance at this wavelength. The initial rate,  $V_0$ , is determined by the slope of the line of absorbance versus time at the interval in which the fastest activity is observed. The activity of DaHMGR begins to slow down as the concentration of the substrate and cofactor decline and become depleted.

In Figure 6, steady-state parameters were determined based on a non-linear regression fitting the Michaelis-Menten equation done in GraphPad Prism 6.0 with respect to NADH and NADPH and HMG-CoA. The  $K_m$  for HMG-CoA using its preferred cofactor, was calculated to be 7.34  $\mu$ M. The  $k_{cat}$  which was calculated by dividing the  $V_{max}$  by the molar enzyme concentration, was determined to be 22.29  $s^{-1}$  and the  $k_{cat}/K_M$ ,  $3.04 \times 10^6$   $M^{-1}s^{-1}$  [12].

With regards to cofactor preference, a clear preference for NADH is observed. The  $K_m$  is determined to be 28.48  $\mu$ M, the  $k_{cat}$  is 16.58 and the  $k_{cat}/K_M$  is  $5.82 \times 10^5$   $M^{-1}s^{-1}$ . While DaHMGR is capable of using NADPH, when NADPH was assessed, the enzyme did not reach  $V_{max}$  at the concentrations that were measurable using the detection limit of the instruments used, showing low activity for NADPH. As observed by a linear relationship between initial rate versus NADPH

concentration in Figure 6, despite increasing the concentration of NADPH as high as 500  $\mu$ M, the enzyme never reaches a maximum velocity. Thus, steady-state parameters are not determined [12].

Table 5 shows a summary of the novel kinetic parameters determined for DaHMGR.

Inhibition by citrate with respect to HMG-CoA was observed when supplementing steady-state kinetics varying HMG-CoA with 0-355 mM Citrate. Double-reciprocal plots with linear fits show no difference in the y-intercepts of the linear fits with and without citrate suggesting competitive inhibition as shown in Figure 6. The apparent  $K_M$  after supplementing the reaction with citrate was increased, and as such a  $K_i$  of 170 mM was calculated.

### 3.3.2. Crystal structures

#### 3.3.2.1. Apo DaHMGR

The initial apo DaHMGR crystallized with one monomer in the asymmetric unit such that its catalytically related adjacent monomer is generated by crystallographic symmetry. Only density for the residues 3-376 was observed and was modeled with the protein chain. The CTD, presumably disordered, make up the remaining residues, 377-429. An overlay of this apo-DaHMGR and the remaining structures presented in this chapter are shown in Figure 7 revealing no significant differences in the overall structure of the main domain which excludes the CTD. The data collection and refinement statistics of this structure are provided in Chapter Two.

#### 3.3.2.2. Mevalonate-bound DaHMGR

Similar to the apo structure above, mevalonate-bound DaHMGR crystallized with one monomer in the asymmetric unit and did not have density for the CTD. Density consistent with mevalonate was observed and was modeled into the structure. A Polder  $mF_o - DF_c$  omit density

map with respect to mevalonate calculated in Phenix[13] is shown in Figure 8, which includes the interactions between mevalonate and the surrounding residues within the active site. An overlay of mevalonate-bound DaHMGR and the remaining structures presented in this chapter are shown in Figure 7 revealing no significant differences in the overall structure of the main domain.

### 3.3.2.3. NADH- and mevalonate-bound DaHMGR

Similar to the above apo and mevalonate-bound structures, NADH- and mevalonate-bound DaHMGR crystallized with one monomer in the asymmetric unit and did not have density for the CTD. Density consistent with NADH and mevalonate was observed and was modeled into the structure. Polder  $mF_o - DF_c$  omit density maps with respect to NADH and mevalonate calculated in Phenix[13] are shown in Figure 9. An overlay of NADH- and mevalonate-DaHMGR and the remaining structures presented in this chapter are shown in Figure 7 revealing no significant differences in the overall structure of the main domain which excludes the CTD.

### 3.3.2.4. NADH- and citrate-bound DaHMGR

DaHMGR, bound to citrate and NADH, crystallized with three molecules in the asymmetric unit. Chains A and B form a homodimer, whereas chain C forms a homodimer with a second chain C monomer (C') through crystallographic symmetry. Density for the CTD was observed in chain B and C, whereas the CTD of chain A was unresolved. The CTD is observed in a similar location as the previously described closed conformation observed in the ternary complex of PmHMGR (PDB 1QAX) however it appears to be turned upside-down. The linker region, which connects the main domain to the CTD, allows the three helices of the CTD to be oriented such that there is a  $\sim 180^\circ$  in-place rotation of the CTD, a conformation we term the “flipped conformation.”

The final model contains the following residues: chain A contains residues 2-377, chain B contains residues 2-428, and residues C contains residues 2-429. Polder  $mF_o - DF_c$  omit density maps with respect to citrate and NADH calculated in Phenix[13] are shown in Figure 10 and Figure 11, respectively, which includes the interactions between citrate and NADH and the surrounding residues in the active site. An overlay of NADH- and citrate-bound DaHMGR and the remaining structures presented in this chapter are shown in Figure 7 revealing no significant differences in the overall structure of the main domain which excludes the CTD. The root-mean-square deviation for deviation for C $\alpha$  atoms (rmsd) of 0.75–0.84 when compared to the structure of cofactor-bound SpHMGR (PDB entry 5WPJ)[12].

#### 3.3.2.5. Citrate-bound DaHMGR

For this experiment, DaHMGR was crystallized with HMG-CoA but there was no electron density for the CoA moiety of HMG-CoA. Instead, density consistent with citrate, which was part of the crystallization reagent, with a final concentration of ~50 mM citrate, was observed, in a similar manner as described previously. This citrate-bound DaHMGR structure crystallized with two monomers in the asymmetric unit that form the catalytic homodimer. Density for residues 3-429 were observed in both chains, however, density for the CTD in Chain A was poorer than the density observed for the CTD in Chain B. Crystal packing analysis suggest that the CTD of Chain B has more interactions with symmetry-related monomers than the CTD of Chain A which may be the reason poorer density is observed for the latter.

The CTD adopts the flipped conformation previously observed in the NADH- and citrate-bound DaHMGR. Notably, the crystallization parameters between the citrate-bound structure and the NADH- and citrate-bound structure, are different. For example, the space group, the respective crystal contacts, the number of molecules observed in the asymmetric unit, all differ between these



two structures, however they both produce structures that have the CTD in a flipped conformation. Figure 12 shows an overlay of CTD of the citrate-bound DaHMGR and NADH- and citrate-bound DaHMGR.

### 3.3.2.6. Citrate- and NADPH-bound SpHMGR

Apo and ligand bound structures of SpHMGR have been previously reported, however a novel crystal morphology in a previously unreported space group was obtained during crystallization screening involving SpHMGR mutants discussed in Chapter Four. Using those conditions, a structure of wildtype SpHMGR co-crystallized with NADPH was determined. This crystal formed the homodimer in the asymmetric unit. In both chains, density for residues 3-424 were observed and modeled in. Density for NADPH and citrate, which was part of the precipitation reagent during crystallization experiments, was observed and modeled in. Polder  $mF_o - DF_c$  omit density maps with respect to NADPH and citrate for one chain were calculated in Phenix are shown in Figure 13.

The CTD adopts a conformation that resembles the “flipped” conformation that was previously observed in citrate- and NADH-bound DaHMGR. The CTD occupies the same region as the closed conformation found in the ternary complex of PmHMGR (PDB 1QAX) but the orientation of the three helices that make up the CTD are flipped. The catalytic histidine, which point towards the active site during distinct steps in the mechanism, is now flipped outwards toward the solvent.

### 3.4. Discussion

HMGR, an NAD(P)-dependent oxidoreductase, is a fascinating enzyme that exemplifies how the structure of these biologically important enzymes accomplish its complex function. Its intrigue stems from its ability to catalyze the four-electron reduction of HMG-CoA to mevalonate, using two cofactors that individually need to enter and leave at specific times during the reaction mechanism. Moreover, the enzyme can control the flux of ligands, from substrate to product, by preventing unnecessary loss of intermediates, thus achieving a highly efficient process.

To date, research has revealed important facets of the active site as well as the role of a flexible flap domain, named the C-terminal domain (CTD), that contributes towards its mechanism of catalysis, which is owed to studies done on several HMGRs that incrementally added to our understanding of how HMGR works. With gaps in our understanding remaining, this research sought to characterize a lesser-studied HMGR, from *Delftia acidovorans* (DaHMGR), with two main goals: 1) to kinetically characterize HMGR with respect to substrate and cofactor and solve crystal structures of its preferred cofactor bound; 2) to solve crystal structures that contain the CTD to substantiate its role in the mechanism. Additional structures of SpHMGR and SpHMGR mutants, which are the focus of Chapter Four, were also consulted as they pertain to the role of the CTD.

Regarding the first goal to kinetically characterize DaHMGR, one specific aim was to quantify the cofactor preference of DaHMGR for either NADH or NADPH. Here, using steady-state kinetics, the evidence shows DaHMGR as an NADH-dependent enzyme, with a  $K_m$  of  $28.48 \pm 1.39 \mu\text{M}$  and  $k_{\text{cat}}$  of  $16.58 \pm 0.24 \text{ s}^{-1}$  for a catalytic efficiency of  $5.82 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ . These kinetic parameters are comparable with respect to various characterized HMGRs. For example,  $K_m$  values for various HMGRs range from 12.9-160  $\mu\text{M}$  as listed on Table 4. Catalytic efficiency is also

within reported ranges as compared to, for example, SpHMGR which has a  $k_{\text{cat}}/K_m$  of  $2.4 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ .

While DaHMGR can utilize NADPH, the enzyme did not reach maximal velocity using the range of NADPH used here, further showing DaHMGR's preferred utility of NADH over NADPH. Another specific aim was to perform steady-state kinetics to reveal the affinity of DaHMGR to HMG-CoA. Here, we present DaHMGR kinetics with respect to HMG-CoA to have a  $K_m$  of  $7.34 \mu\text{M}$  and  $k_{\text{cat}}$  of  $22.29 \pm 0.48 \text{ s}^{-1}$  for a catalytic efficiency of  $3.04 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$  when using its preferred cofactor. Of the class II HMGRs that have been investigated, DaHMGR has the lowest  $K_m$ , suggesting a relatively high affinity for DaHMGR to HMG-CoA.

With the knowledge that DaHMGR prefers NAD(H), we solved the crystal structure of DaHMGR with NADH bound. We confirm an important motif that was previously hypothesized to directly contribute towards cofactor specificity. It was previously proposed that a short  $\alpha$ -helix found only in class II HMGR termed the "cofactor helix" plays a crucial role in determining cofactor specificity [3]. At least three important residues in positions one, three and seven of this helix are involved in governing cofactor specificity. In NADH-preferring HMGR, for example in PmHMGR, these three residues are made up of an aspartate that is conserved among NADH-preferring HMGRs and two hydrophobic leucine residues. The two hydrophobic residues in the third and seventh position stack against the aromatic adenine ring. More importantly, the aspartate sidechain hydrogen interacts with the 2-hydroxyl of the NAD ribose.

This aspartate likely contributes towards selectivity by preventing the 2'-phosphate of its non-preferred cofactor, NADPH, from binding. In contrast, for NADPH-preferring HMGRs, like in SpHMGR, a tyrosine is observed instead of aspartate of the cofactor helix and residues serine

and arginine, occupy positions three and seven. Instead of nonpolar residues found in NADH-preferring HMGRs, serine and arginine making hydrogen bonding and pi-pi stacking interactions with the ribose phosphate group and adenine respectively. A summary of the cofactor helix as a structural motif that confers cofactor specificity is as follows: DX(V/L)XXXL for NADH and YXSXXXR for NADPH, where “X” indicates variable amino acids. In the crystal structure presented here, DaHMGR contains the cofactor helix sequence DKVLIGL. Figure 14 shows the cofactor helices that constitute the developing model used to explain cofactor specificity, which alongside the kinetic data obtained here, agrees with previous hypothesis of the important role in dictating cofactor specificity of the cofactor helix. Further work into the cofactor specificity, and engineering efforts are explored in the next chapter.

A second goal of this section in further characterizing DaHMGR and elucidating the mechanism of catalysis among class II HMGRs was to solve crystal structures with resolved CTDs. In many structures of HMGRs, there is no density for this domain because it is a highly flexible and dynamic domain. HMGR is believed to be active in its crystal form [8], so the flexibility of the CTD in the context of its native function persists in the crystal lattice, thus making it difficult to procure crystal structures of HMGR that have resolved CTDs. However, in this study, we were successful in obtaining several structures of DaHMGR and SpHMGR and SpHMGR mutants that reveal novel locations of the CTD.

In one structure of NADH-bound DaHMGR which contained citrate in the active site – citrate was part of the crystallization reagents –we reveal a surprising new orientation of the CTD of class II HMGR. Prior to this discovery of this new orientation, the CTD was previously observed in either an open, closed, or partially closed conformation, depending on what ligands were bound. In the structure of apo-SpHMGR with no ligands bound, the CTD assumed an open conformation,

allowing entry of substrate and cofactor. In the structure of with NADPH-bound SpHMGR, the CTD was in a similar open conformation, which suggests that cofactor binding alone does not trigger closing of the CTD. In the structure of HMG-CoA-bound SpHMGR, the CTD was located in a partially closed conformation where the CTD interacts with CoA moiety of the substrate, closing off the substrate-binding site while leaving the cofactor-binding site exposed. When HMG-CoA and NAD<sup>+</sup> were bound simultaneously to PmHMGR, a closed conformation where the CTD covers the active site was observed. The novel location found in this study shows DaHMGR CTD having a flipped conformation which further expands this paradigm.

It is worth pointing out that crystal packing interactions appear to have allowed the capture of the CTD in this flipped conformation whereby symmetry-related monomers stack up against the CTD to stabilize it. This observation may question the physiological relevance of the CTD in this conformation. However, it is important to consider that the crystal packing interactions that contribute towards stabilizing the CTD in two different structures, namely the citrate-bound DaHMGR structure and the NADH- and citrate-bound structure, are different. In NADH- and citrate-bound DaHMGR, the crystal packing interactions that stabilize the CTD are primarily made up of the CTD of one monomer and the CTD from a symmetry-related monomer. However, in citrate-bound DaHMGR, the CTD from one monomer is stacked up against two CTDs from two adjacent symmetry-related monomers in a vastly different orientation. Moreover, in citrate-bound DaHMGR, the way the CTD of chain A stacks up against two CTDs from the adjacent monomer is different from the way chain B stacks up against two CTDs from its adjacent monomer. The difference between the crystal packing interactions in both the above structures are depicted in Figure 15.

As noted in the Results, the two above structures are from entirely different crystal lattices, and crystallized in different space groups, with different crystal packing, yet still reveal the CTD in a flipped conformation. Even within two monomers of the same crystal structure (citrate-bound DaHMGR), the crystal packing interactions that stabilize each CTD differ, despite the conformation of the CTD remaining the same. This suggests that the flipped conformation of the CTD may not simply be a crystallographic artifact. The observation that two completely different crystal packing configurations captured the CTD in this conformation is evidence of this conformation being intrinsic to the protein's native function. Furthermore, as will be highlighted later in this discussion, this flipped conformation is similarly observed in a different citrate-bound HMGR homolog.

According to the CTD movements previously observed, the CTD of NADH-bound DaHMGR would be expected to assume an open conformation just as with NADPH-bound SpHMGR. However, in the flipped conformation, the NADH-binding site and the substrate-binding site are covered by the CTD, with Arg405 and Arg428 of the CTD interacting with the cofactor. In this conformation, the catalytic histidine side chain is flipped out of the active site and is pointing towards solvent. In addition, while the location of the three helical bundle of the CTD is in the same general vicinity of the closed structure as observed in PmHMGR, the orientation of the three helices is different than previously observed. In the closed structure, the first helix of the CTD stacks up against the active site, with the catalytically important histidine pointing towards the active site, as it lays on the surface of the opposite monomer. The two remaining helices are positioned away from the active site on the outer surface, pointing towards the solvent. In the flipped conformation of DaHMGR, the first helix is now on the outer surface of the protein, and instead, the two subsequent helices lay on the surface of the protein. Figure 16 shows the location

and orientation of the flipped conformation in comparison to the closed conformation observed in PmHMGR.

What is reiterated by this observation is the inherent flexibility of the CTD and that the CTD may assume multiple conformations that facilitate the progression of the reaction. In the flipped conformation, crystal packing interactions have captured a snapshot of one possible location that the CTD can assume when cofactor is bound. Thus, it is likely that there is no single conformation for cofactor-bound HMGR but rather conformations that point to the inherent mobility of this domain upon cofactor binding. The degree of flexibility of this domain is owed to a linker region that connects the main domain of the protein to the mobile CTD. The linker region is found preceding the three helical bundle that comprises the CTD and extends from a helix from the main domain to the CTD. While occupying the same overall location, the linker region reveals how it can afford the CTD's ability to adopt various conformations and locations as shown in Figure 17, including the novel flipped location revealed in this study. Interestingly, in the ternary complex of NAD<sup>+</sup>- and HMG-CoA-bound PmHMGR, where the CTD adopts a closed conformation, this linker region forms part of a bent helix that connects the main domain to the CTD. It is possible that the winding and unwinding of this helix contributes towards the flexibility of this linker region and subsequently the positioning of the CTD.

In addition to highlighting the flexibility of the CTD, this flipped location may also provide greater insight into the catalytic mechanism. Upon analysis of citrate bound to the active site, it was proposed that citrate, which resembles the structures of intermediates mevaldehyde and mevalonate, may be mimicking an intermediate step. As previously mentioned, NADH-bound DaHMGR would be expected to have its CTD in an open conformation, allowing entry and exit of both cofactor and substrate. However, in NADH-bound DaHMGR, both sites are closed off,

thus suggesting an intermediary step where the CTD is preventing the release of an intermediate, blocking off the binding site of the substrate/product. To further support the idea that citrate may be mimicking an intermediate, we observe many of the same interactions between citrate and DaHMGR that were also observed in a prior structure of mevalonate-bound PmHMGR[10]. As such, the flipped conformation may in fact represent an intermediate mevaldehyde- or mevalonate-bound structure.

It is important, therefore, to examine the resemblance of citrate to the intermediates. Mevaldehyde contains a terminal carbonyl at a sp<sup>2</sup>-hybridized carbon, while mevalonate contains a primary alcohol at a sp<sup>3</sup>-hybridized carbon. Figure 10 shows a comparison between the structures of citrate and intermediates. Citrate, with a terminal sp<sup>2</sup>-hybridized carboxylate carbon, arguably resembles mevaldehyde more closely, and if so, the citrate-bound DaHMGR presented here may provide useful insight into the catalytic mechanism. Citrate was also found to inhibit the reaction competitively with respect to HMG-CoA as shown in Figure 6.

Additionally, the flipped conformation confirms that while His381 plays an important role in catalysis, its direct role is not during HMG-CoA reduction or mevaldehyde reduction but rather in the protonation of CoA as it is cleaved. In this scenario, His381 would not need to be in the vicinity of the active site in the subsequent steps when mevaldehyde is reduced to mevalonate using the second cofactor. The citrate-bound DaHMGR indeed has His381 in the flipped conformation located towards the solvent.

Incidentally, while pursuing mutant constructs of SpHMGR related to cofactor preference efforts which will be discussed in the next Chapter, crystal structures revealed locations of the CTD that appear to corroborate some of the assertions made here. Two SpHMGR chimeric



proteins, SpHMGR containing a DaHMGR cofactor helix with and without M185T mutations, both bound with their newly preferred cofactor, NADH, reveal new locations of the CTD. The CTD in both of these structures are in the general open location. This agrees with the above observation that there is no single conformation and location of the CTD when only cofactor is bound. Figure 18 shows the open conformations of these newly acquired structures overlaid with a previous structure of SpHMGR in an open conformation.

An additional structure, of SpHMGR bound by NADPH and citrate, reveals the CTD in an apparent flipped conformation. In this structure, the location of the CTD is in the general vicinity of the previously observed closed conformation of PmHMGR, however, the orientation of the helical bundle more resembles the flipped conformation of DaHMGR. In this flipped conformation, the catalytically important histidine is indeed pointing away from the active site and is solvent exposed. Similarly observed in the flipped CTD conformation in DaHMGR, the first helix points towards the solvent, whereas the remaining two helices lay closer to the surface of the protein. Figure 19 shows the comparison between the flipped conformations between DaHMGR and SpHMGR. In both DaHMGR and SpHMGR flipped conformations, citrate is bound, and as proposed above, may be a mechanistically relevant conformation in possibly mimicking an intermediate step in the mechanism.

Taken together, the results presented thus far in this Chapter provide new insight into the steps of catalysis and as outlined in Figure 20, the result is a proposed revised mechanism [12]. In apo form, without any ligands bound, the CTD is flexible and disordered which is corroborated by the apo-DaHMGR structure presented here as well as other apo-HMGRs structures. There are many open conformations, but common among these, is the ability for substrate and cofactor to bind. It is not known whether substrate or cofactor binds first and there are structures of only

HMG-CoA bound and structures of only NADP(H) bound. Upon cofactor binding, however, it is evident that the CTD remains flexible as confirmed by the several open conformations of NADPH-bound SpHMGR and the NADH-bound DaHMGR structures presented here. If, instead, HMG-CoA binds first, the CTD can adopt a partially closed conformation as observed in HMG-CoA-bound SpHMGR which shows how HMG-CoA can be retained while allowing NADP(H) to bind[12].

Once the ternary complex is formed, the CTD assumes a closed conformation as depicted in NAD<sup>+</sup>- and HMG-CoA-bound PmHMGR. In this closed conformation, the CTD situates the catalytic histidine (His381 in PmHMGR numbering) at the active site. As illustrated in Figure 20 of the mechanism, HMG-CoA is reduced to mevaldyl-CoA utilizing the first equivalent of cofactor and it has been proposed that either Glu83 or Lys267 protonates the substrate carbonyl. Structures of HMG-CoA-, and the mevalonate and citrate-bound structures presented here, all show conserved positioning of these two residues as they interact with the oxygen atom of the carbonyl [12].

Further mutational analysis of the roles of the above two catalytic residues, Glu83 and Lys267, alongside structural analysis offer insight into the roles of these two residues. Mutation of Lys267 to alanine significantly diminishes activity, but does not abolish it completely [10]. In contrast, mutation of Glu83 to glutamine, essentially kills activity, and computational analysis has suggested that transfer of a hydride from the cofactor is more favorable if Glu83 is protonated [6], [14]. The revised proposed mechanism, therefore, suggests that Glu83 acts as the catalytic acid, protonating the substrate carbonyl, whereas Lys267 acts to polarize the carbonyl, priming it for the first reduction step. It is worth noting that this revised mechanism does not rule out Lys267 as

being the source of protonation, but the structural and functional data thus far informs this proposal [12].

Following this first reduction step, and once the intermediate mevaldyl-CoA is formed, it is uncertain the order of the immediate steps that follow. Specifically, it is unclear the order of steps as they pertain to the release of the used oxidized cofactor and the incorporation of the second reduced cofactor, and the steps involved in the cleavage of mevaldyl-CoA to form mevaldehyde and CoA. With evidence suggesting that a CoA-containing intermediate can remain bound to the enzyme, and the observation that in the partially closed conformation of SpHMGR showing how the CTD binds the CoA moiety while keeping the cofactor binding site open for exchange gives credence to the presumption that the release of the oxidized cofactor and the acquisition of a reduced cofactor, occurs while mevaldyl-CoA remains bound [3], [8]. Subsequently, this intermediate is decomposed to mevaldehyde and CoA [12].

In the decomposition of mevaldyl-CoA to mevaldehyde and CoA step, Glu83 may be involved in deprotonating mevaldyl-CoA. From the citrate-bound DaHMGR and SpHMGR structures presented here, which possibly mimic a mevaldehyde-bound HMGR, indeed Glu83 is shown interacting with the corresponding oxygen atom of citrate. This suggests that Glu83 may be protonated when this oxygen is deprotonated. Following cleavage of CoA, His381 from the CTD of the opposite monomer likely protonates the thiolate moiety, which is then able to be released, leaving behind mevaldehyde in the active site. In the citrate-bound flipped conformation structure of DaHMGR, it is noted that the CTD would sterically clash with any bound CoA. As shown in Figure 21, which is an overlay of HMG-CoA-bound PmHMGR and the citrate-bound DaHMGR structure, the location of the CTD sterically clashes with CoA and would contradict

CoA still being bound. As such, this flipped conformation suggests that CoA can be released once mevaldehyde is formed and may not remain bound for the remainder of the reaction [12].

In the citrate-bound SpHMGR structure, the flipped CTD conformation does not appear to sterically hinder CoA if CoA were bound. An overlay of HMG-CoA-bound SpHMGR and citrate-bound SpHMGR reveals that CoA can still be present. The steric clashes shown above in the case of the flipped CTD conformation of DaHMGR are not observed in the flipped conformation of SpHMGR. In this case, it is possible that CoA can be retained and released at a later step in the mechanism. However, while the citrate-bound SpHMGR conformation of the CTD includes the possibility that CoA can be released later in the mechanism, based on the flipped structure of DaHMGR, CoA is likely released once mevaldehyde is formed.

At this point in the mechanism, the second cofactor and mevaldehyde are bound, therefore one would expect that the enzyme can undergo a conformation change to adopt the closed conformation of the CTD which would enable the reduction of mevaldehyde to mevalonate. In the citrate-bound DaHMGR mimicking mevaldehyde binding, the mevalonate bound-DaHMGR, as well as prior mevalonate-bound PmHMGR structures, show Glu83 or Lys267 interacting with the analogous oxygen atom. This suggests that one of these two residues protonates mevaldehyde as it forms mevalonate alcohol. As discussed above, regarding the roles of Glu83 and Lys267 in protonating a prior intermediate, we propose Glu83 as the protonating residue. Following this step, the CTD can become flexible and mobile, and can adopt an open conformation, which would allow the release of mevalonate and oxidized cofactor. This final step leaves the active site empty and exposed to the solvent for reprotonation of active site residues, and ultimately, ready for substrate and cofactor to bind for a new round of the mechanism[12].

In this chapter, the biochemical and crystallographic data presented here illuminate the structural and functional dynamics of the class II HMGR using DaHMGR and SpHMGR homologs. Using DaHMGR, by solving its crystal structure bound to its preferred NADH cofactor, we have confirmed the importance of a helical motif that appears to determine the cofactor preference of the enzyme. Furthermore, the various crystal structures of apo, mevalonate-bound, NADH- and citrate-bound DaHMGR, and citrate-bound DaHMGR as well as citrate-bound SpHMGR, provides a snapshot into the flexibility of the CTD, adopting a novel flipped location that contributes towards and expands the proposed HMGR mechanism. With citrate bound, this research may also provide insight into the intermediary steps of the reaction cycle. Collectively, this research makes considerable progress towards the hope of growing our understanding of the catalytic mechanism and cofactor specificity of this biologically important enzyme.

### **3.5. Future work**

#### **3.5.1. Additional kinetic characterization**

While this study makes significant strides in characterizing the kinetic behavior of DaHMGR, additional functional studies are warranted. Steady-state kinetics to expand the Michaelis-Menten parameters of DaHMGR with respect to the oxidized cofactors as well as mevalonate and CoA are examples of work that needs to be explored. Since the  $K_M$  of HMG-CoA is relatively low compared to other HMGRs, suggesting that the affinity of the substrate is high in the forward reaction converting HMG-CoA to mevalonate, it would be worthwhile to determine if DaHMGR also binds the reverse direction substrates with similar affinity. Also, as will be discussed in Chapter Five, it is likely that DaHMGR prefers to operate in the reverse direction,

thus steady-state kinetics to characterize its affinity to the oxidized form of NADH will be insightful.

### 3.5.2. Additional structures with different ligands bound

The new work presented here unearths exciting directions for further characterization of DaHMGR and therefore class II HMGR, including the hopes of obtaining additional crystal structures of various ligands bound – either intermediates, substrate, or products, with or without cofactors. This will further the insight into the structural dynamics required for catalysis. It will be advantageous to have a library of structures that correspond to various points in the mechanism, ideally the conformations of the CTD resolved. These structures include DaHMGR bound to: HMG-CoA + NAD<sup>+</sup>, mevalonate + CoA + NADH, and NADP(H). Especially in light of the flipped conformation with citrate possibly mimicking the mevaldehyde intermediate and its implications on the mechanism, it would be desirable to solve the structure of DaHMGR or SpHMGR with an intermediate bound, for example, having a structure of DaHMGR+mevaldehyde. Such a structure would be helpful in discerning the physiological relevance of the flipped conformation.

### 3.5.3. Understanding the forces that enable CTD motions

As the CTD traverses through the various locations, which so far have been depicted in open, partially closed, flipped, and closed conformations, there are instances where residues of the CTD interact with moieties of the rest of the protein complex. To further interrogate the importance of such conformations, it would be worthwhile to pursue mutational analysis that perturbs these interactions, to observe how catalytic activity is influenced. If these various locations are mechanistically important, then mutating residues that are involved in the interaction between the

CTD and the rest of the protein might diminish activity, and such an analysis would home in on the significance of the various conformations and implicated residues.

Similarly, further interrogation into the role of the linker region that connects the CTD to the rest of the protein would be helpful in determining its part in accommodating or even promoting the ability of the CTD to adopt various locations. Based on Figure 17, in the partially closed conformation of the CTD, the linker region is comprised partly of a helix. In the closed conformation, the linker region is almost entirely helical. Whereas in the other conformations, the linker region does not form a distinct secondary structure. Thus, it appears that the winding and unwinding of this linker region may contribute towards the opening and closing of the CTD. One could pursue investigation into this phenomenon by, for example, disrupting helical formation by introducing a residue such as proline, and observe the consequences on catalysis.

### **3.6. Materials and methods**

DaHMGR protein expression and purification are outlined in the Materials and Methods of Chapter Two. In this section, the Materials and Methods associated with DaHMGR kinetic and structural analysis are detailed. In addition, as mentioned above, the crystal structures of various SpHMGR mutant constructs, which are the focus of Chapter Four, contain novel insight into the conformations of the CTD and therefore will be included in the Discussion of this chapter, however the Materials and Methods for those structures are outlined in Chapter Four.

#### *3.6.1. In vitro* enzyme activity assays

Steady-state kinetics to acquire Michaelis-Menten parameters,  $K_m$  and  $k_{cat}$ , were conducted by monitoring the enzyme-catalyzed oxidation of NAD(P)H via a decrease in absorbance at 340

nm, using an extinction coefficient of  $6200 \text{ M}^{-1} \text{ cm}^{-1}$ . For assessing cofactor parameters, HMG-CoA was kept constant at saturating concentrations while varying the concentration of NAD(P)H, and for assessing substrate parameters, NAD(P)H was kept at saturating concentrations while varying HMG-CoA concentration.  $K_m$  was calculated using a non-linear regression fit of the graph of initial rate versus substrate concentration to the Michaelis-Menten equation in GraphPad Prism 7[15], and  $k_{cat}$  was determined by dividing  $V_{max}$  by the molar enzyme concentration. Experiments were conducted in triplicate using a microplate reader. The reaction mixtures were prepared as shown Table 3, with the reaction being initiated by the addition of HMG-CoA.

### 3.6.2. Citrate inhibition assays

To assess inhibition with respect to HMG-CoA, the above steady-state kinetics with the same reaction mixtures when assessing HMG-CoA parameters were conducted by including 0, 188, and 350 mM Sodium Citrate (pH 7.4). Double reciprocal plots were used to analysis the nature of inhibition and to determine the  $K_i$ .

### 3.6.3. Crystallization

#### 3.6.3.1. NADH- and citrate- bound DaHMGR

To attempt crystallization of DaHMGR, sparse-matrix screening was performed using a Crystal Gryphon crystallization robot (Art Robbins) with ~500 different crystallization conditions by sitting-drop vapor diffusion. Three different enzyme mixtures using DaHMGR at 8.5 mg/mL were prepared for this screen: DaHMGR alone, DaHMGR with 1 mM NADH, and DaHMGR with 1 mM HMG-CoA + 1 mM NAD<sup>+</sup>. Crystals were identified in many conditions, and using Izit dye, these crystals were believed to be protein crystals, however a majority developed an irregular shape not suitable for X-ray diffraction (see Figure 22). Condition 63 of Crystal Screen HT (Hampton



Research) which contains 0.5 M ammonium sulfate, 0.1 M sodium citrate tribasic dihydrate pH 5.6, and 1.0 M lithium sulfate monohydrate showed promising crystals and subsequent optimization of this condition was performed. The initial crystals were flat semi-circles that grew into overlapping clusters. Optimization of these conditions involved varying the pH (from pH 5 to pH 7) of citrate buffer and varying concentrations of lithium sulfate (200 mM to 1200 mM) using hanging-drop vapor diffusion at room temperature. Optimal three-dimensional hexagonal crystals were obtained at 0.1 M sodium citrate pH 5.0, 500 mM ammonium sulfate and 4-800 mM lithium sulfate.

In preparation for X-ray diffraction, the optimized hexagonal crystals were cryoprotected using 0.5 M ammonium sulfate, 0.1 M sodium citrate tribasic dehydrate pH 5.0, 600 mM lithium sulfate, 26% glycerol or ethylene glycol supplemented with the same concentrations of ligands in the co-crystallization before being flash-cooled in liquid nitrogen. As discussed in the results, this yielded the first crystal structure of DaHMGR, with citrate and NADH bound.

#### 3.6.3.2. Apo DaHMGR, mevalonate-bound DaHMGR, and NADH- and mevalonate-bound DaHMGR, and citrate-bound DaHMGR

Following the acquisition of citrate- and NADH-bound DaHMGR, we sought to acquire additional structures of DaHMGR. As alluded to in the Materials and Methods section of Chapter Two, this was done by modifying the crystallization condition with other buffers, adjusting pH, and including other precipitating reagents. By modifying the crystallization buffer to Bis-Tris pH 5.5-6.5 and adding polyethylene glycol 3350 ~18-19% we were able to acquire crystals of apo and mevalonate-bound DaHMGR that were suitable for diffraction experiments (See Figure 22).

The apo structure was described in Chapter Two. 1 mM mevalonate was added in a similar condition as apo –100 mM Bis-Tris (pH 6.5), 600 mM lithium sulfate, and 19% PEG 3350 –to obtain the mevalonate-bound structure. These crystals were soaked in a cryoprotectant solution composed of the crystallization solution with 1 mM mevalonate, 1 mM NAD<sup>+</sup>, and 21% sucrose before being flash-cooled in liquid nitrogen to acquire the NADH- and MEV-bound structures.

During subsequent high-throughput screens using DaHMGR, a novel crystal morphology was obtained. These crystals, which looked like cubes, were crystallized with 1 mM HMG-CoA, and were looped directly from the screens without optimization for crystallization experiments. The condition belonged condition B2 of Wizard Cryo Screen from Hampton Research and contained 100 mM Sodium phosphate dibasic/Citric acid pH 4.2, 200 mM Ammonium Sulfate, 10% (v/v) glycerol and 20% (v/v) PEG 300. As will be discussed, HMG-CoA was not bound, instead citrate was observed at the HMG-CoA binding site, similarly observed in the NADH- and citrate-bound DaHMGR above. This citrate-bound DaHMGR, however, crystallized in both a different morphology and space group than all prior structures.

#### 3.6.3.3. Citrate- and NADPH-bound SpHMGR

During crystallization screening exploring cofactor specificity, outlined in Chapter Four, crystals of SpHMGR mutants that formed in a previously unreported space group were found. Utilizing similar conditions, SpHMGR was co-crystallized with 10 mM NADPH in 100 mM citrate pH 5.0, 17% PEG 3350 and yielded hexagonal-shaped crystals. Crystals were looped without cryoprotectant, and flash-frozen for data collection.

#### 3.6.3.4. X-ray Data Collection, structure determination, and refinement

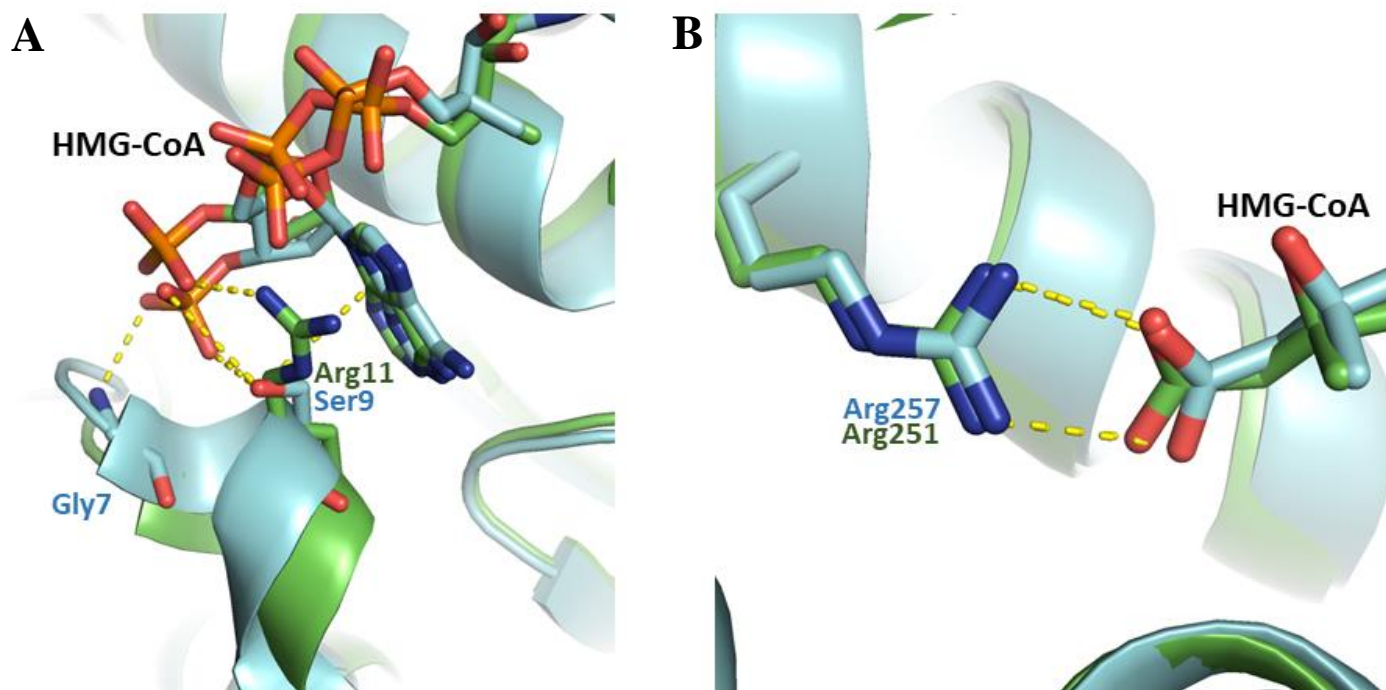
X-ray diffraction data were collected at the Advanced Photon Source (APS). For the DaHMGR with both citrate and NADH bound, the data were indexed, merged, and scaled using MOSFLM[16] in space group  $P6_52_12$  with three molecules in the asymmetric unit at 2.14 Å resolution. Using the PHENIX [13] suite, Phaser [17] was used to determine the structure by molecular replacement using the structure of PmHMGR (PDB entry 1QAX) with the substrate, cofactor and C-terminal domain removed as the search model. Electron density for the C-terminal domain was observed in only chains B and C. Using Coot [18], a poly-alanine sequence was built first, prior to refinement in Phenix, and sequence-specific residues subsequently modelled in. Electron density for NADH was also observed; Phenix.elbow [19] was used to generate NADH restraint files that was used in the refinement of NADH in the structure. Electron density consistent with citrate, which was part of the crystallization buffer, was found in the active site and modelled in. Phenix.refine was used to perform reciprocal space refinement, with model-building in Coot.

X-ray data collection, structure determination and refinement were done similarly for the apo, citrate- and mevalonate-bound structures. For these structures, however, instead of PmHMGR, DaHMGR from the above NADH- and citrate-bound DaHMGR was used as the search model for DaHMGR crystals, with the ligands removed, when the space groups were different from the above. Apo, and mevalonate-bound crystals belonged to space group  $P3_22_1$  and the citrate-bound crystal belonged to space group  $P2_13$ . The same test set of reflections was preserved for both apo and mevalonate-bound DaHMGR data sets. In addition, no density for the CTDs were observed in the initial apo and mevalonate-bound structures, however, the citrate-bound DaHMGR structure contained electron density for the CTD which was modeled as described above. For SpHMGR, apo SpHMGR (PDB 3QAU) was used as a search model. These crystals formed in  $P3_22_1$  with two molecules in the asymmetric unit. Density for citrate, found in the crystallization

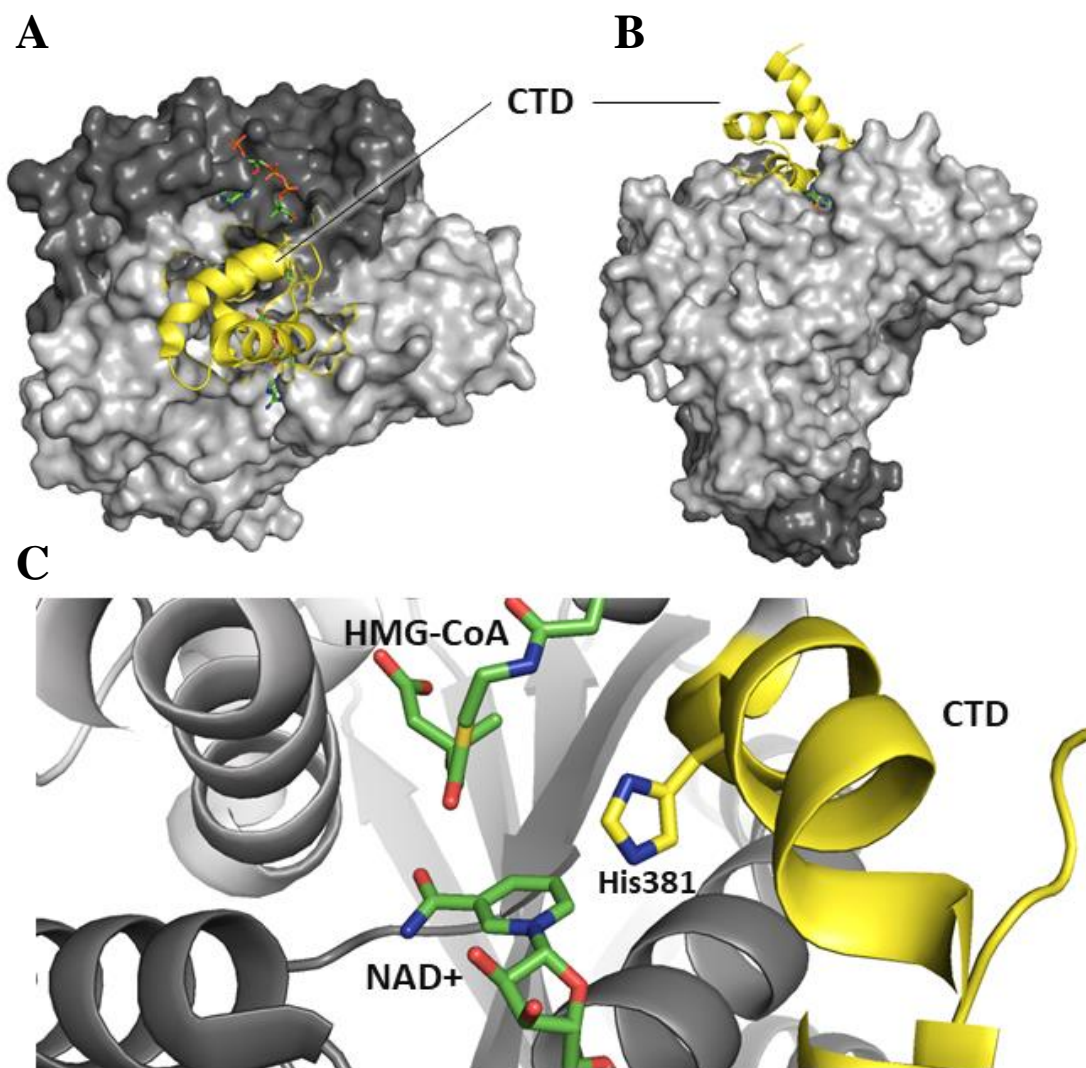
reagent, and NADPH was observed and modeled in as well as for the CTD in both chains. Final data collection and refinement statistics for published structures are listed on Table 1 and unpublished structures are listed on Table 2.





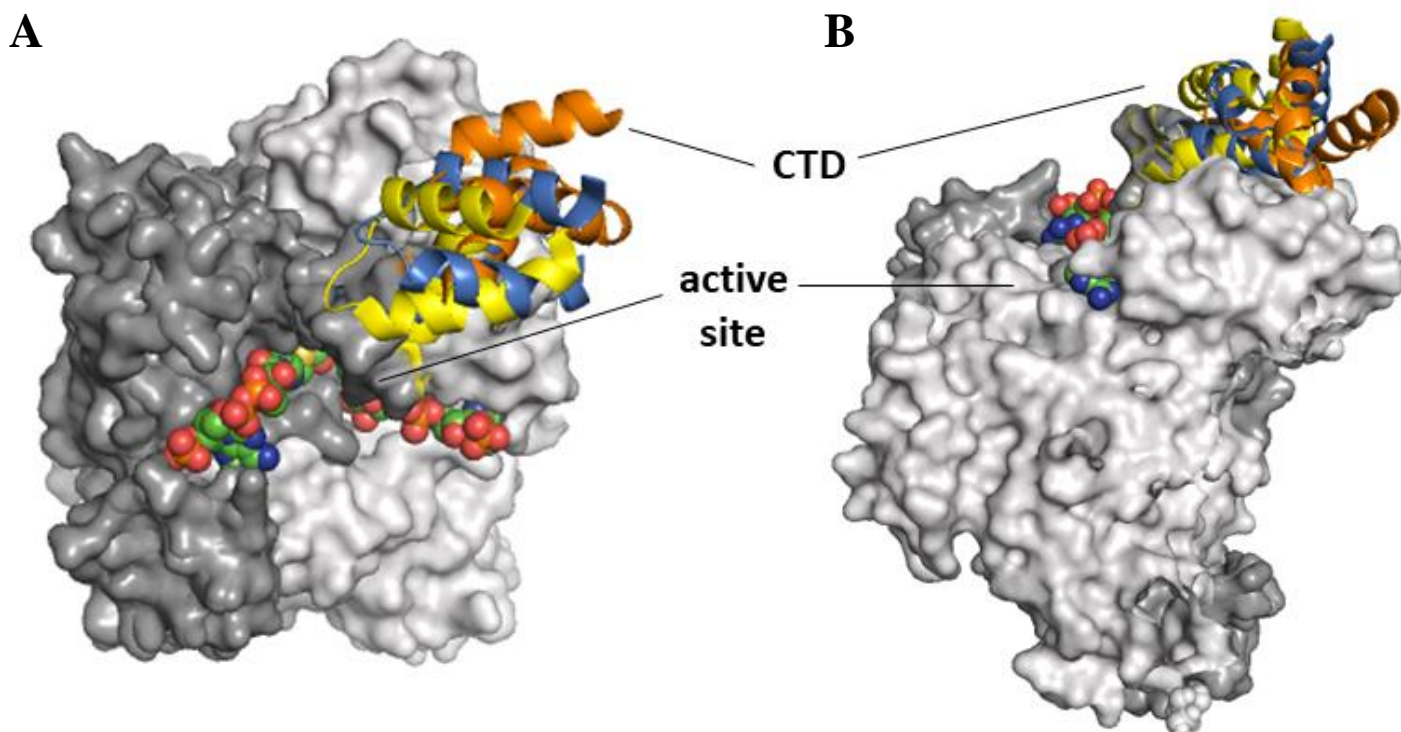


**Figure 2. Overlay of HMG-CoA binding pocket.** PmHMGR (PDB 4I6A) shown in green and SpHMGR (PDB 5WPK) shown in cyan in cartoon representation and HMG-CoA in sticks. Coloring of atoms are as follows: C colored green or cyan in PmHMGR or SpHMGR respectively, N colored blue, O colored red, P colored orange, and S colored. **A)** adenine ring, solvent exposed end of HMG-CoA and **B)** HMG moiety bound at the interior.

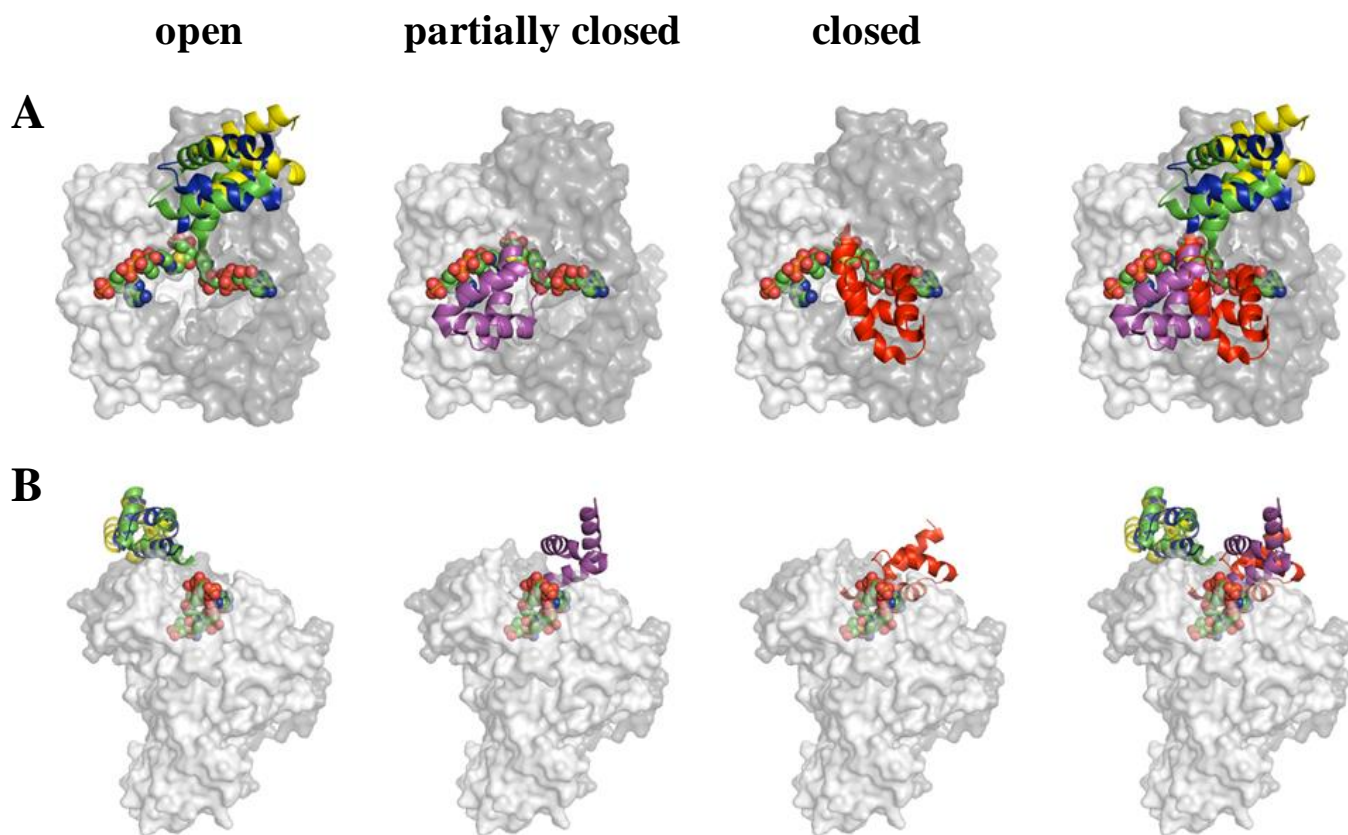


**Figure 3. Closed conformation of the C-terminal domain.** PmHMGR bound by NAD<sup>+</sup> and HMG-CoA (PDB 1QAX) in a ternary complex, revealing the location of the CTD in a closed conformation. Each monomer of the homodimer is colored a shade of gray and shown in surface representation, with the CTD of one monomer, shown in yellow cartoon representation, covering the active site. NAD<sup>+</sup> and HMG-CoA are shown in sticks with the following coloring: C colored green, N colored blue, O colored red, P colored orange, and S colored. The catalytically important histidine, His381 from the CTD, is depicted in stick model, showing the vicinity of this residue to the active site as a result of the CTD adopting this closed conformation. **A)** top view of the CTD covering the active site. **B)** a side view of the CTD with respect to the rest of the complex. **C)** Close up of the active site.

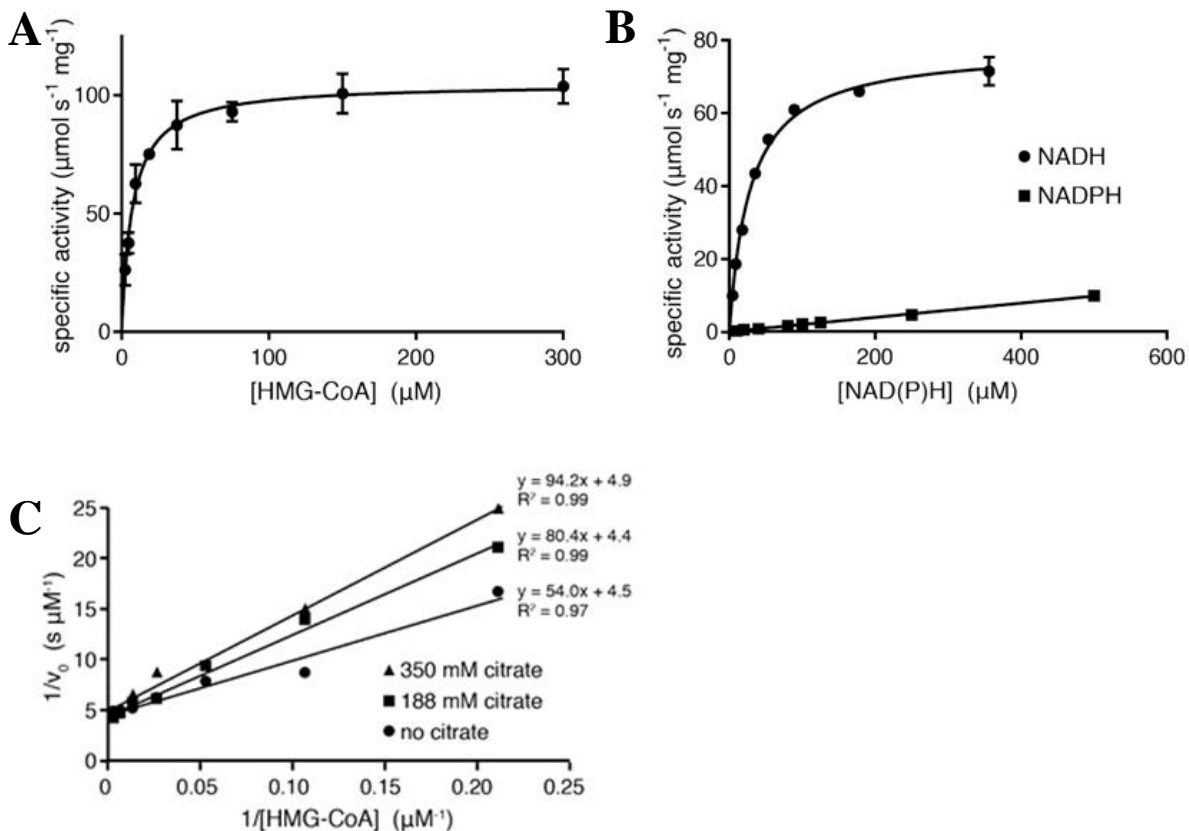




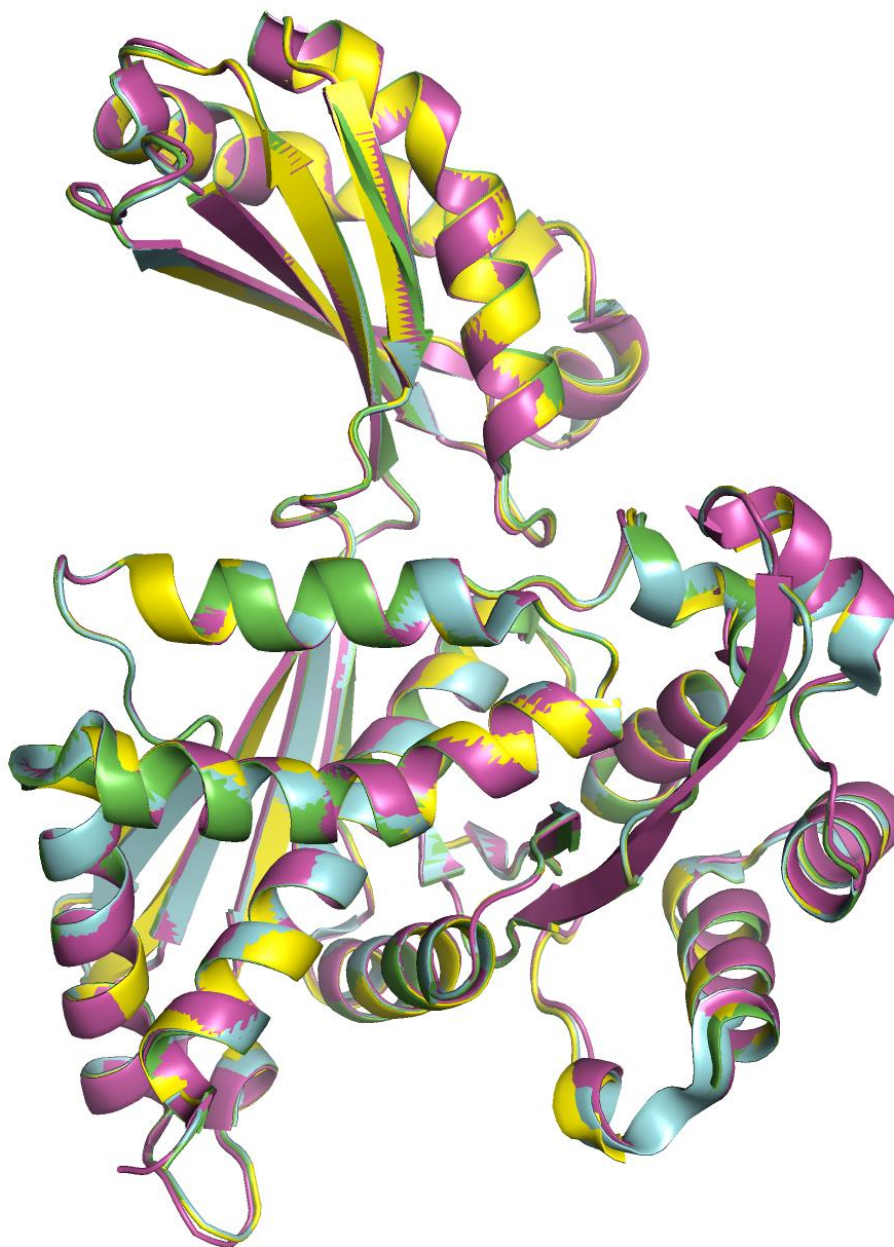
**Figure 4. Open conformations of the C-terminal domain found in various SpHMGR structures (PDB 3QAU and PDB 5WPJ).** Each monomer of the homodimer is colored a shade of gray and shown in surface representation, with the CTD in open conformations as observed in SpHMGR structures shown in cartoon representation and colored differently. HMG-CoA and NADP are modelled in from PmHMGR (PDB 1QAX) and SpHMGR (PDB 5WPJ) respectively, to visualize the location of the active site with respect to the open conformation of the CTD. Ligands are colored using the following: C colored green, N colored blue, O colored red, P colored orange, and S colored. **A)** top view of the active site and **B)** a side view.



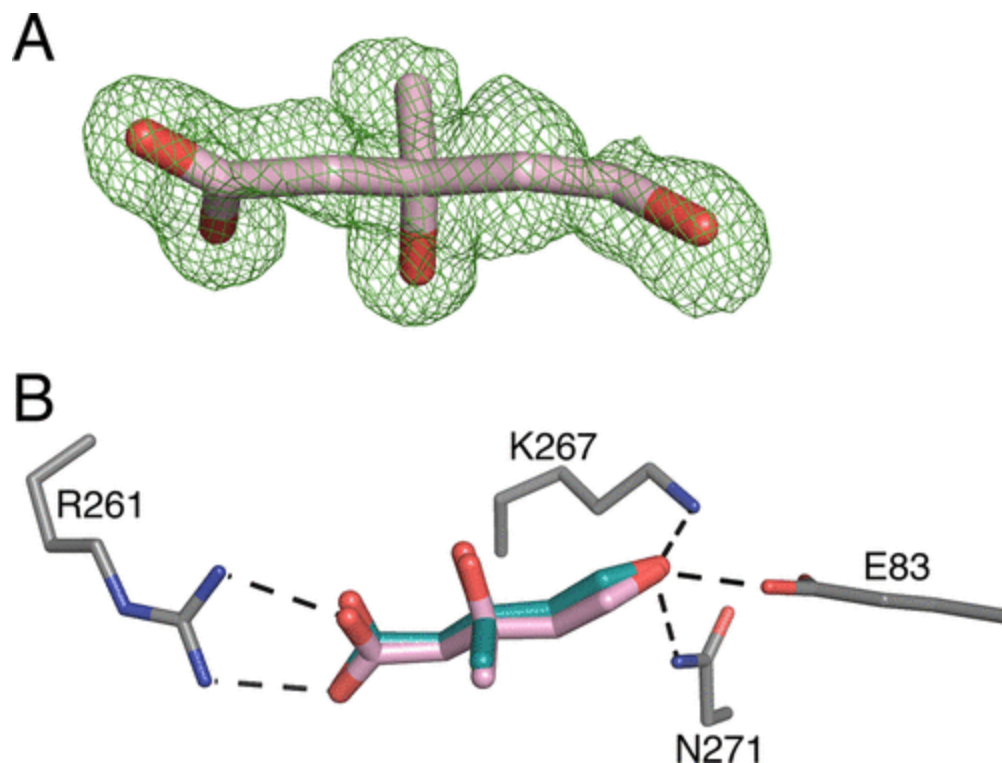
**Figure 5. Various conformations of the C-terminal domain of PmHMGR and SpHMGR observed prior to the work done in this study.** Each monomer of the homodimer is colored a shade of gray and shown in surface representation. . HMG-CoA and NADP are modelled in from PmHMGR (PDB 1QAX) and SpHMGR (PDB 5WPJ) respectively, to visualize the location of the active site with respect to the open conformation of the CTD. Ligands are colored using the following: C colored green, N colored blue, O colored red, P colored orange, and S colored purple. **A)** top view of the active site and **B)** a side view. Open conformation of the CTD is made up of apo-SpHMGR structures (PDB 5WPJ and 3QAU). Partially closed structure conformation of the CTD is made up of the HMG-CoA-bound SpHMGR (PDB 5WPK). The closed conformation of the CTD is made up of NAD<sup>+</sup>- and HMG-CoA-bound PmHMGR (PDB 1QAX). On the right, an overlay of all the locations of the CTD of PmHMGR and SpHMGR. **A)** top view and **B)** side view.



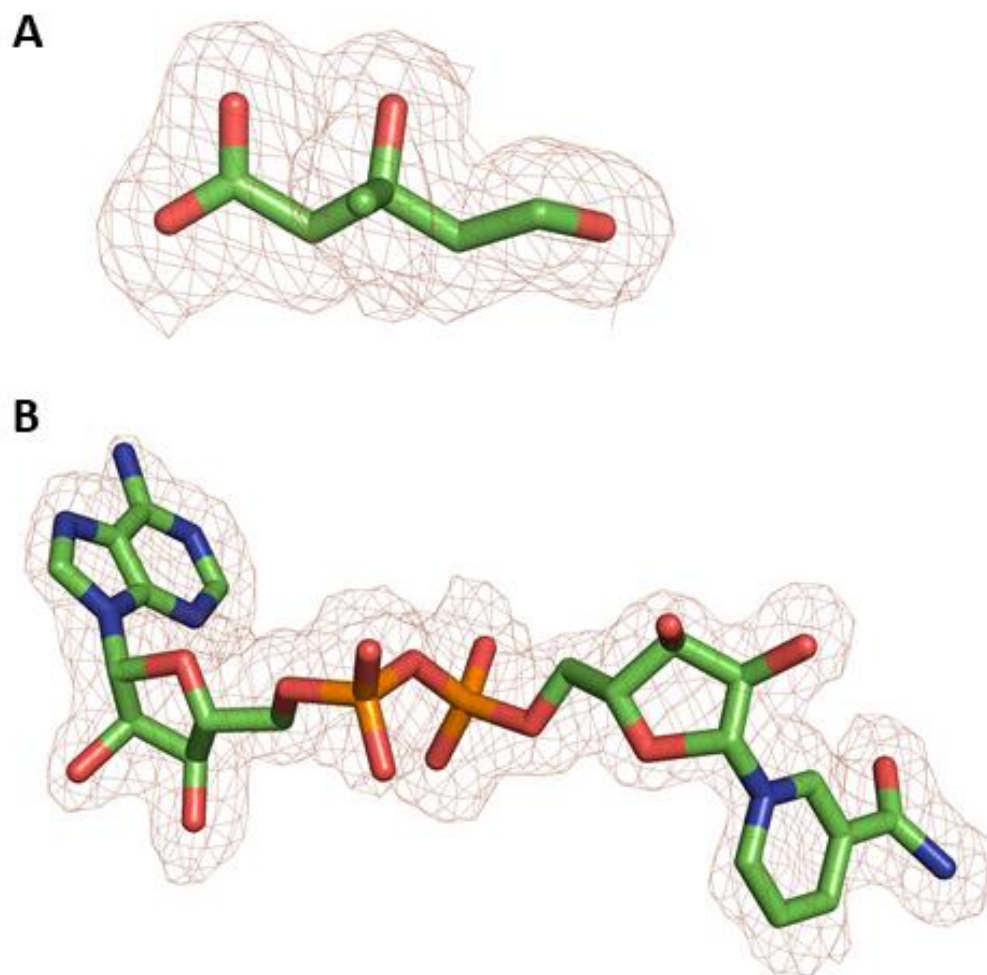
**Figure 6. DaHMGR steady-state kinetics with respect to HMG-CoA, NADH and NADPH, and citrate inhibition.** A) Measuring the specific activity versus HMG-CoA concentration. Concentrations of HMG-CoA ranging from 2 to 300  $\mu\text{M}$  were used. B) Circles represent data with NADH, and squares for NADPH. Concentrations of NADH ranging from 4 to 360  $\mu\text{M}$  NADH were used. Concentrations of NADPH ranging from 5-500  $\mu\text{M}$  were used, but the enzyme did not reach  $V_{\text{max}}$ . Higher concentrations of NADPH could not be assessed due to its high absorbance exceeding the detection limit of the instruments used. C) Competitive citrate inhibition with respect to HMG-CoA with Double-reciprocal plots with linear fits showing citrate competitively binds with respect to HMG-CoA.



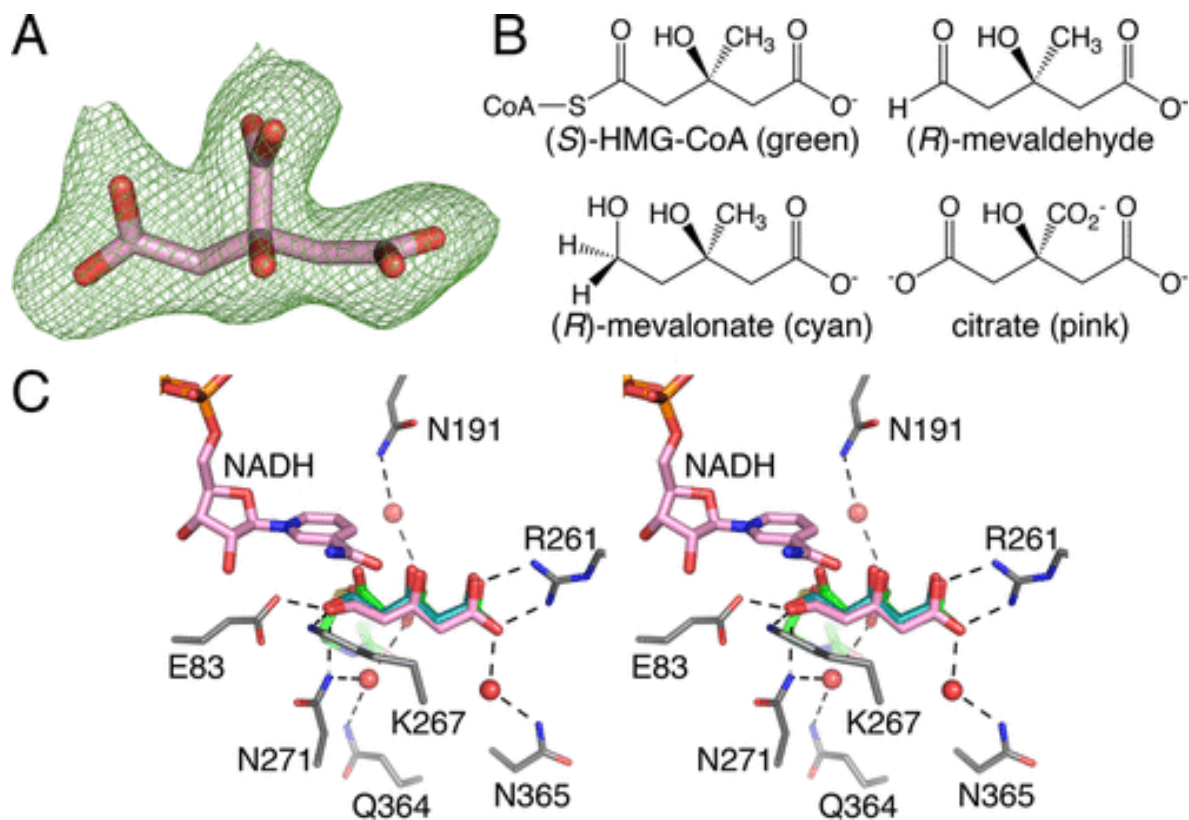
**Figure 7. Overlay of all the DaHMGR structures presented in this chapter.** Cartoon representation of all the DaHMGR structures presented in this chapter, excluding the C-terminal domain, which is only present in the NAD- and citrate-bound DaHMGR structure. Green = apo-DaHMGR, blue = mevalonate-bound DaHMGR, magenta = NAD- and citrate-bound DaHMGR, and yellow = NAD- and mevalonate-bound DaHMGR.



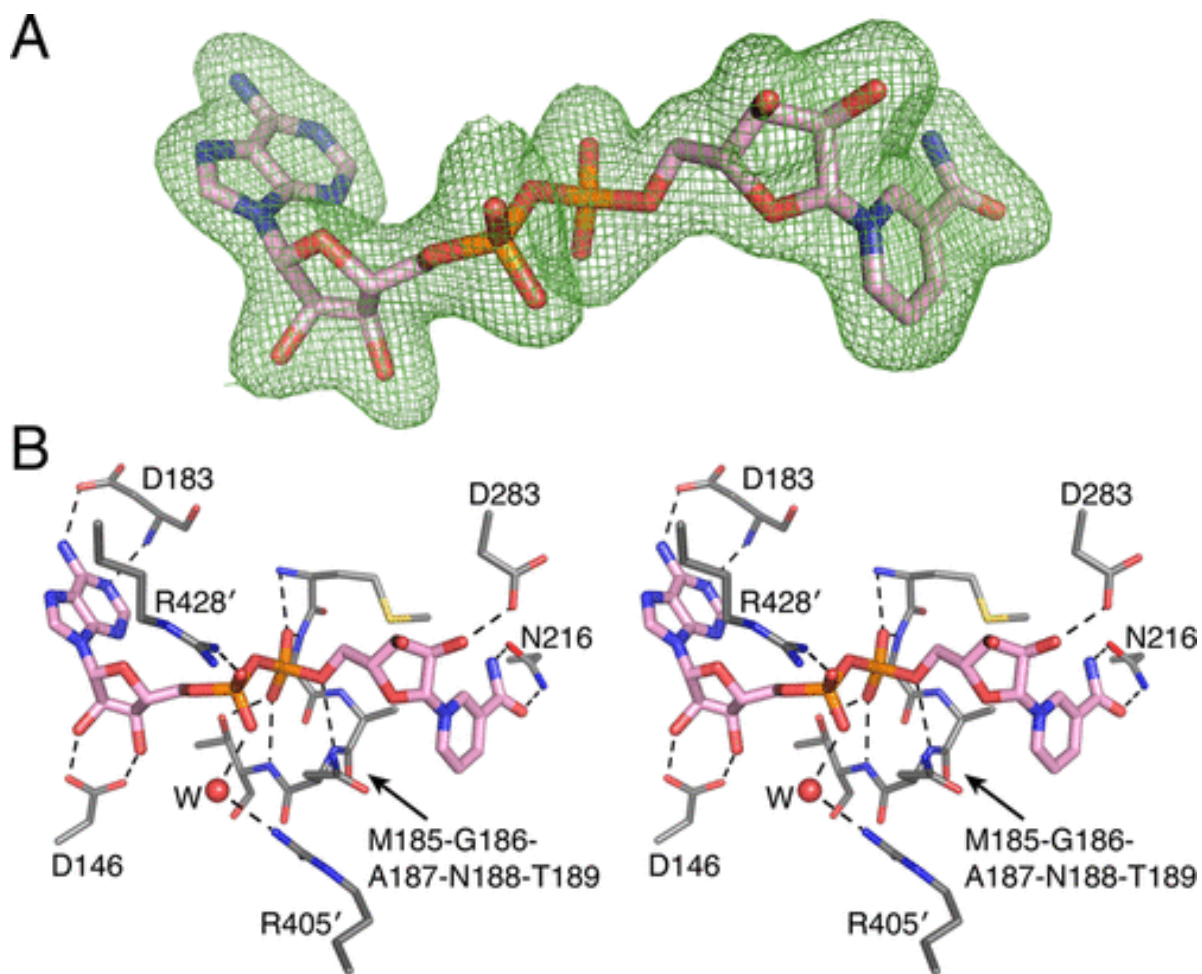
**Figure 8. Mevalonate-binding site of DaHMGR.** **A)** Polder  $mFo - DFc$  omit density contoured at  $3.0\sigma$  shown surrounding mevalonate in green mesh, with mevalonate shown as sticks (C colored pink and O colored red). **B)** The close-up of the substrate-binding site, with mevalonate from DaHMGR (C colored pink), mevalonate from superimposed PmHMGR (PDB entry 1I6Y, C colored teal), and DaHMGR residues (C colored gray) shown as sticks, with N colored blue and O colored red. Dashed lines represent hydrogen bonding or salt-bridge interactions between mevalonate and DaHMGR.



**Figure 9. Omit maps generated with respect to in NAD- and mev-bound DaHMGR.** Polder *mFo* – *DFc* omit density contoured at  $3.0\sigma$  and colored brown surrounding **A)** mevalonate and **B)** NAD with the ligands shown in sticks colored as follows: C colored green, N colored blue, O colored red, and P colored orange.

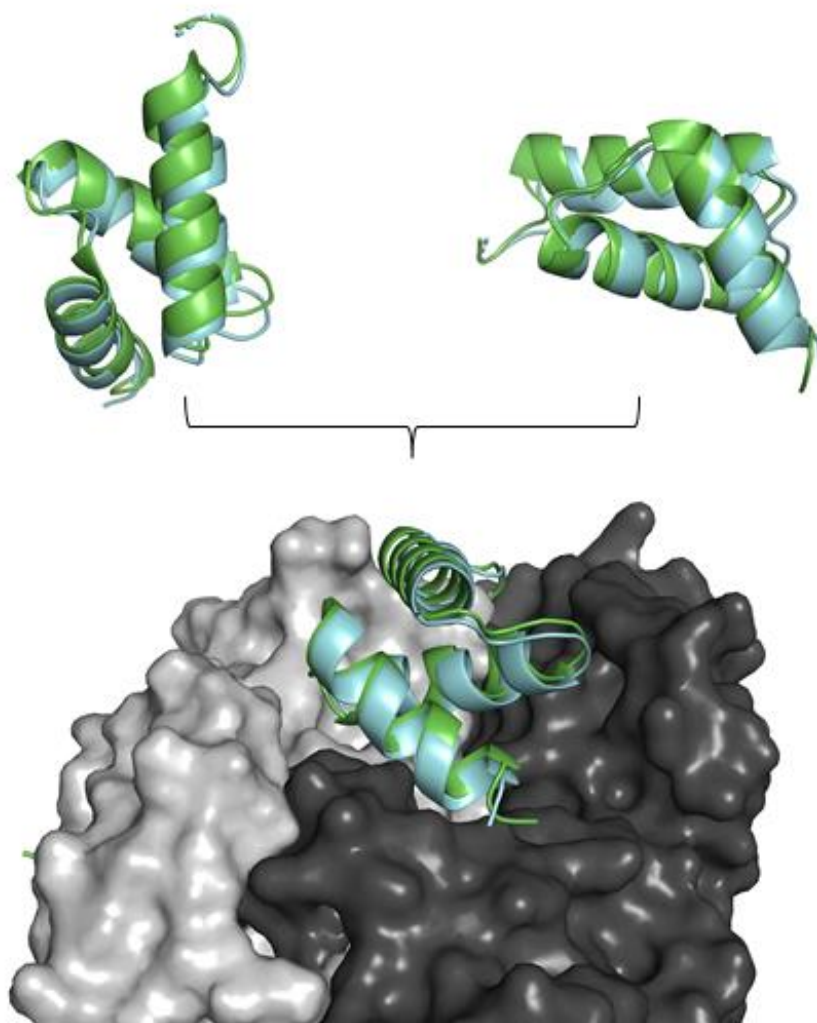


**Figure 10. Substrate-binding site, showing citrate bound in DaHMGR.** **A)** Polder  $mF_o - DFC$  omit density contoured at  $3.0\sigma$  surrounding a citrate molecule shown in green mesh, with citrate shown as sticks. **B)** Comparison between the structures of the substrate (S)-HMG-CoA, intermediate (R)-mevaldehyde, product (R)-mevalonate, and citrate. Annotated colors in parenthesis refer to the corresponding ligand coloring in panel C. **C)** Stereo view of citrate binding in the structure of citrate- and NAD-bound DaHMGR superimposed with mevalonate in the structure of mevalonate-bound DaHMGR and HMG-CoA in the structure of HMG-CoA-bound SpHMGR (PDB 5WPK). Citrate (C colored pink), NADH (labeled, C colored pink), mevalonate (C colored cyan), HMG-CoA (C colored green), and DaHMGR residues (C colored gray) are shown as sticks, with N colored blue, O colored red, P colored orange, and S colored yellow. Dashed lines represent hydrogen bonding or salt-bridge interactions involving citrate and DaHMGR.

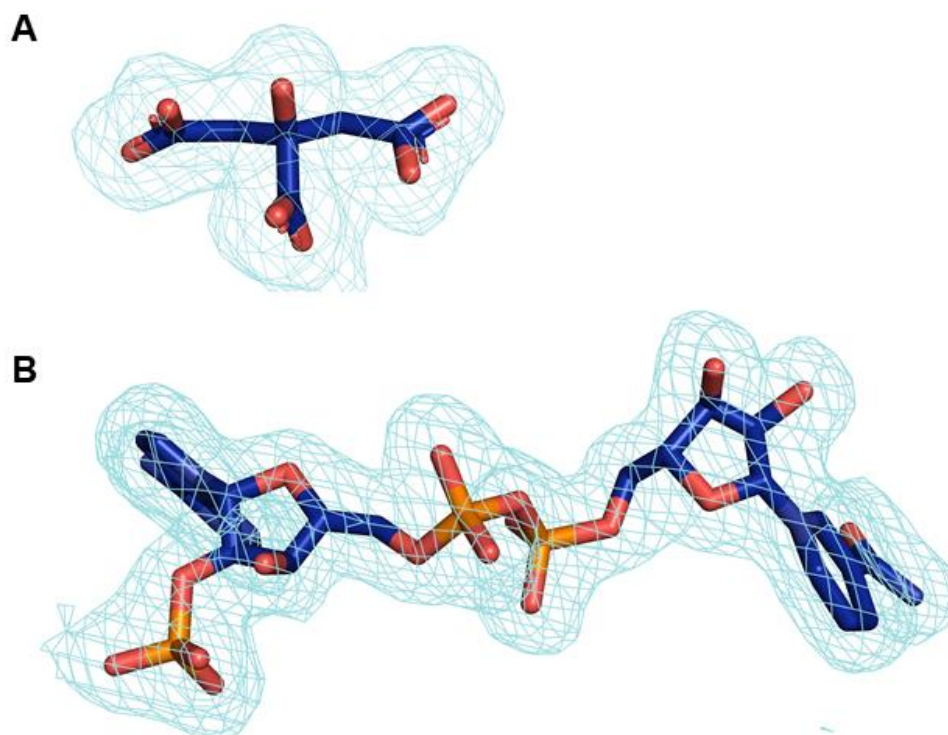


**Figure 11. NAD-binding site.** **A)** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  surrounding NAD shown in green mesh, with NADH shown as sticks. **B)** Stereo view with NADH shown as sticks (C colored pink) and DaHMGR residues shown as sticks (C colored gray), with N colored blue, O colored red, P colored orange, and S colored yellow. A water molecule is shown as a red sphere and labeled “W”. Dashed lines represent hydrogen bonding or salt-bridge interactions.

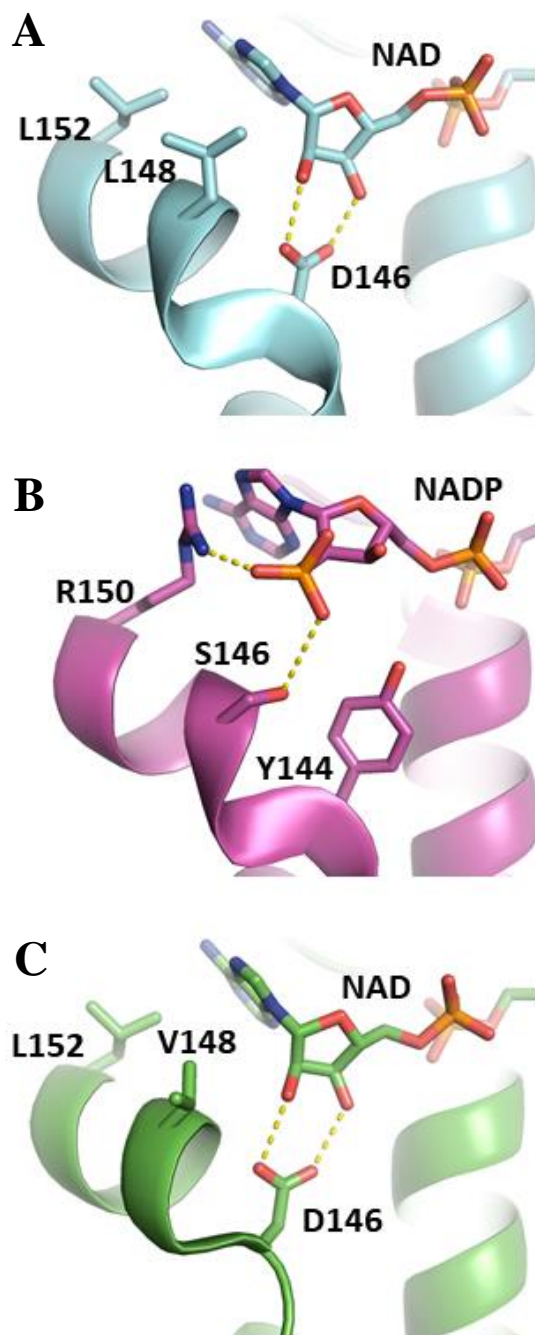




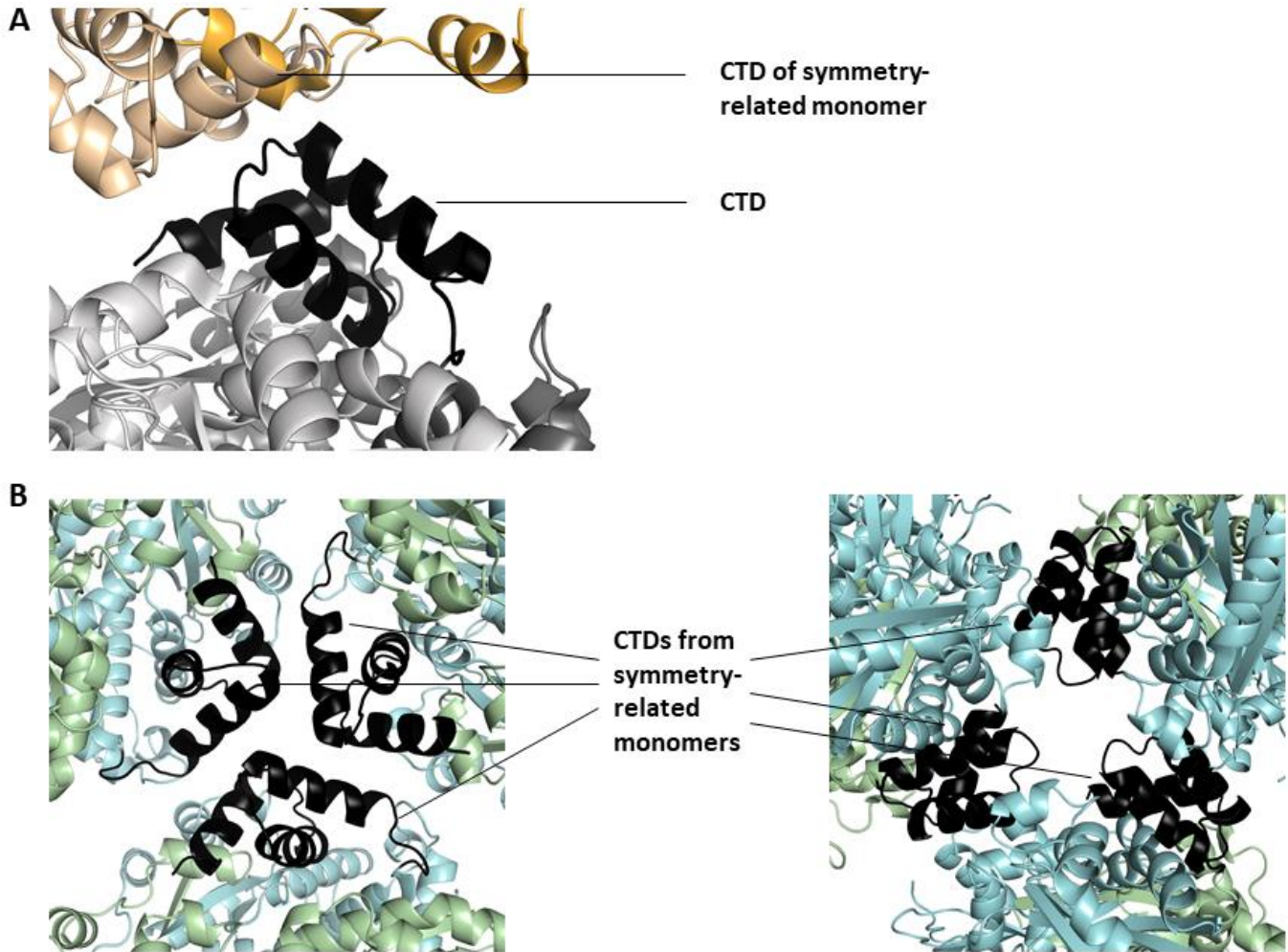
**Figure 12. Comparing the flipped conformation of the C-terminal domain (CTD) observed in two different structures of DaHMGR.** In cartoon representation is an overlay of the CTD of NAD- and citrate-bound DaHMGR shown in green and the CTD of citrate-bound DaHMGR shown in cyan. In surface representation is the rest of the protein where each monomer of the homodimer is colored in a different shade of gray. On top, are two different views of the overlay of the CTDs showing how well they align.



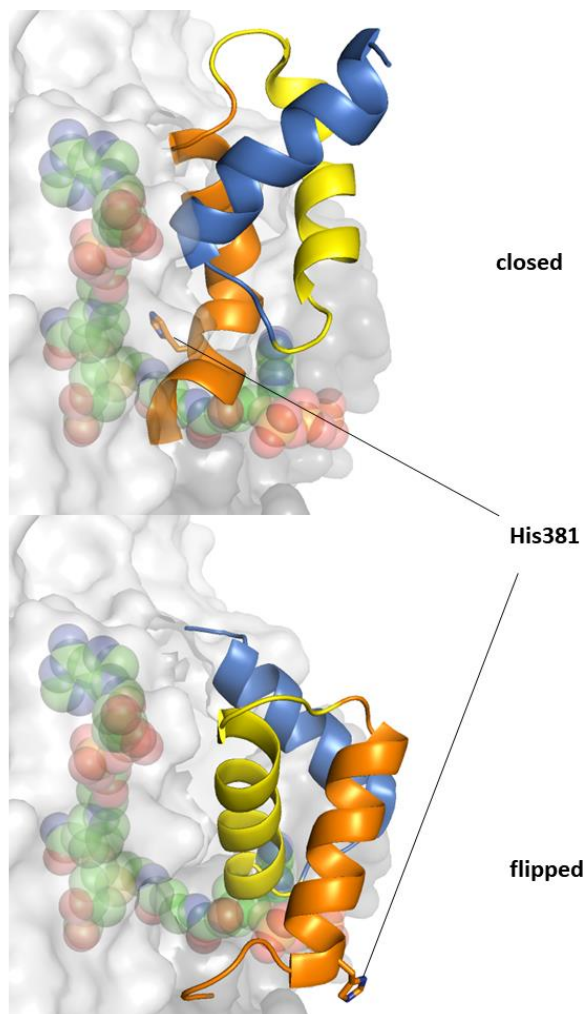
**Figure 13. Polder omit maps generated with respect to citrate and NADPH bound by SpHMGR.** Polder *mFo* – *DFc* omit density contoured at  $3.0\sigma$  and colored brown surrounding **A)** citrate **B)** NADPH with the ligands shown in sticks colored as follows: C colored dark blue, N colored blue, O colored red, and P colored orange.



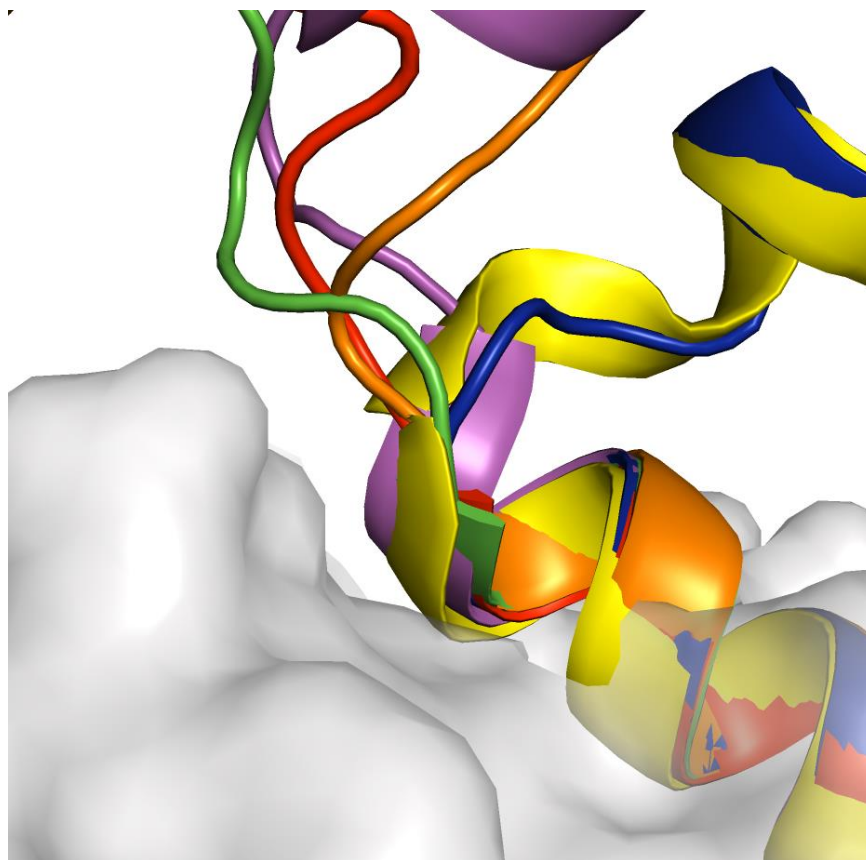
**Figure 14. The cofactor helix that dictates cofactor specific among class II HMGRs. The cofactor helices are shown in cartoon representation. Ligands and important residue are colored using the following: C colored according to the rest of the molecule, N colored blue, O colored red, and P colored orange. A) NAD-bound PmHMGR, B) NADP-bound SpHMGR, and C) NAD-bound DaHMGR.**



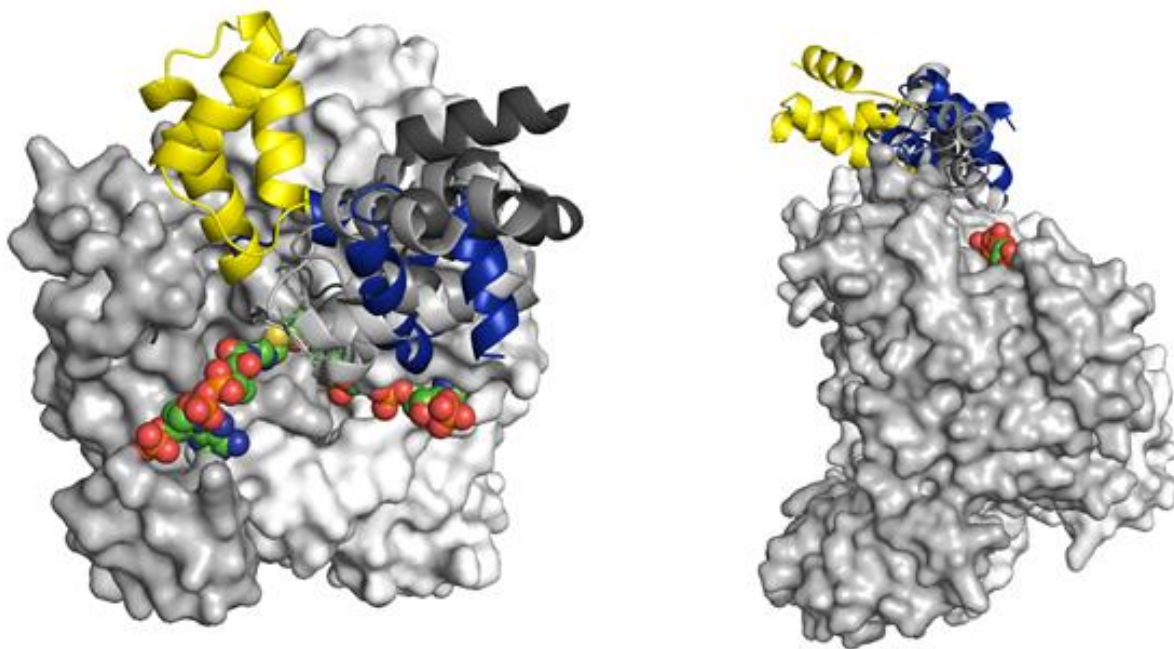
**Figure 15. Different crystal packing that stabilizes C-terminal domain in various DaHMGR structures that adopt the flipped conformation.** Crystal structures are depicted in cartoon representation illustrating the differences in crystal packing that facilitate the stabilization of the CTD in the flipped conformation observed in various structures. **A)** In the NAD- and citrate-bound DaHMGR structure, the CTD of one monomer stacks up against a CTD from a symmetry-related monomer. In gray is the homodimer with the CTD of one monomer in black. In shades of orange are the monomers of a symmetry-related homodimer. The CTD, shown in black, interacts with a single adjacent monomer. **B)** In the citrate-bound DaHMGR structure, symmetry-related monomers stack up against each other such that the CTDs of three monomers interact. On the left is the crystal packing of the CTD of one chain and on the right is from another chain. Green and blue represent a homodimer, and in black are their respective CTDs. This depicts different interactions that stabilize the CTD between the different chains.



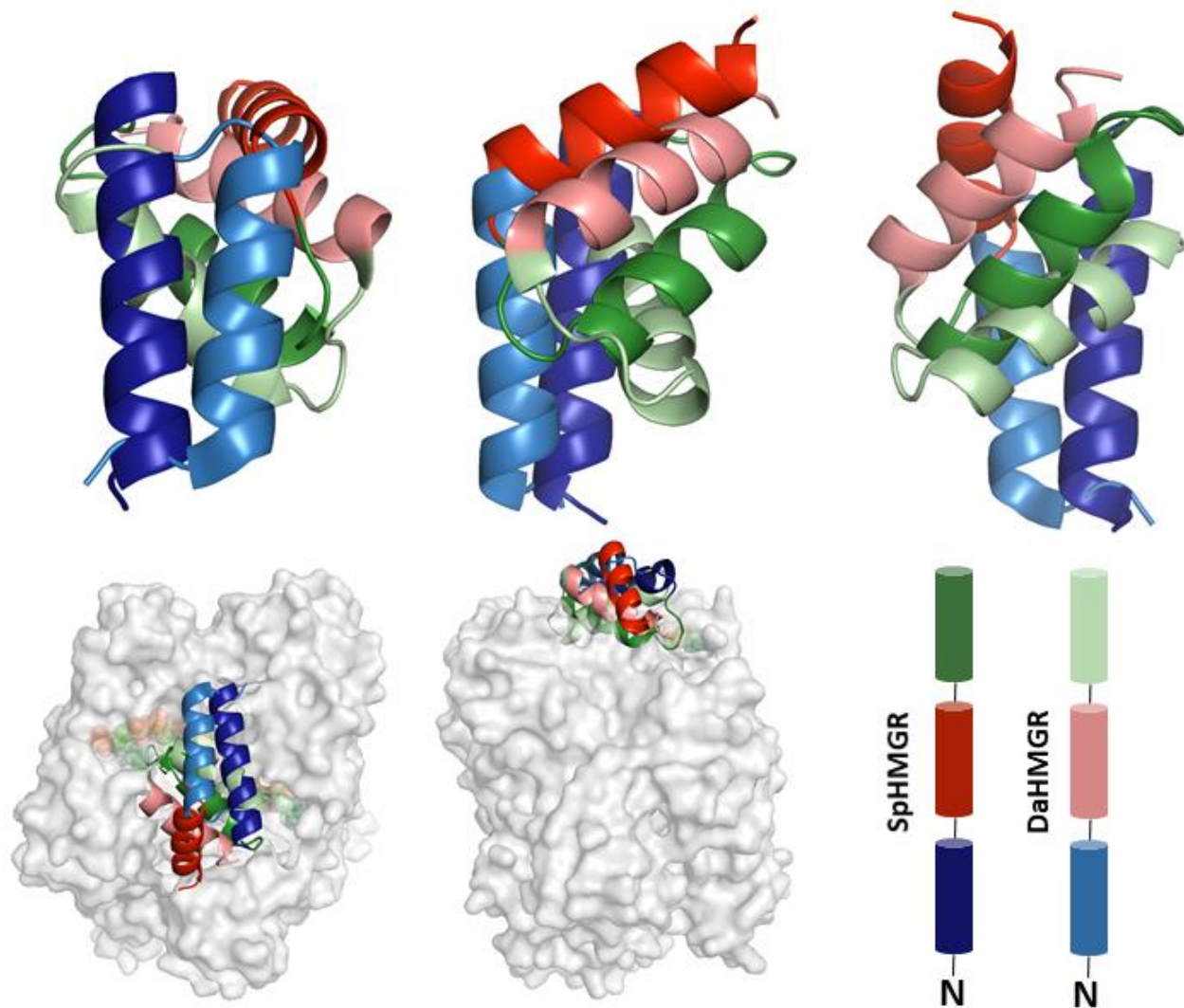
**Figure 16. The C-terminal domain in the closed conformation versus the flipped conformation as observed in PmHMGR and DaHMGR respectively.** Surface representation of the protein homodimer in gray together with cartoon representation of the CTD in the closed conformation of NAD<sup>+</sup>- and HMG-CoA-bound PmHMGR (PDB 1QAX) and the flipped conformation of the CTD in NAD-bound DaHMGR. Each helix of the CTD is colored a different color as follows: first helix = orange, second helix = yellow, and blue helix = third helix. The catalytic histidine, His381 PmHMGR numbering is shown at the active site in the closed conformation versus solvent-exposed in the flipped conformation. Ligands are modeled in from the ternary complex of PmHMGR (PDB 1QAX) and are shown in spheres with the following coloring: C = green, N colored blue, O colored red, S = yellow. and P colored orange



**Figure 17. Linker region, that connects the main body of class II HMGR to the C-terminal domain.** In gray, surface representation of the main body of class II HMGR is depicted, with cartoon representation of the region that links the main body to the CTD. The various colors represent various structures as follows: Green = apo-SpHMGR, (PDB 3QAU). red and green = NADP-bound SpHMGR (PDB 5WPJ), blue = HMG-CoA-bound SpHMGR (PDB 5WPK), yellow = NAD<sup>+</sup>- and HMG-CoA-bound PmHMGR (PDB 1QAX), magenta = citrate- and NAD-bound DaHMGR (PDB 6DIO).

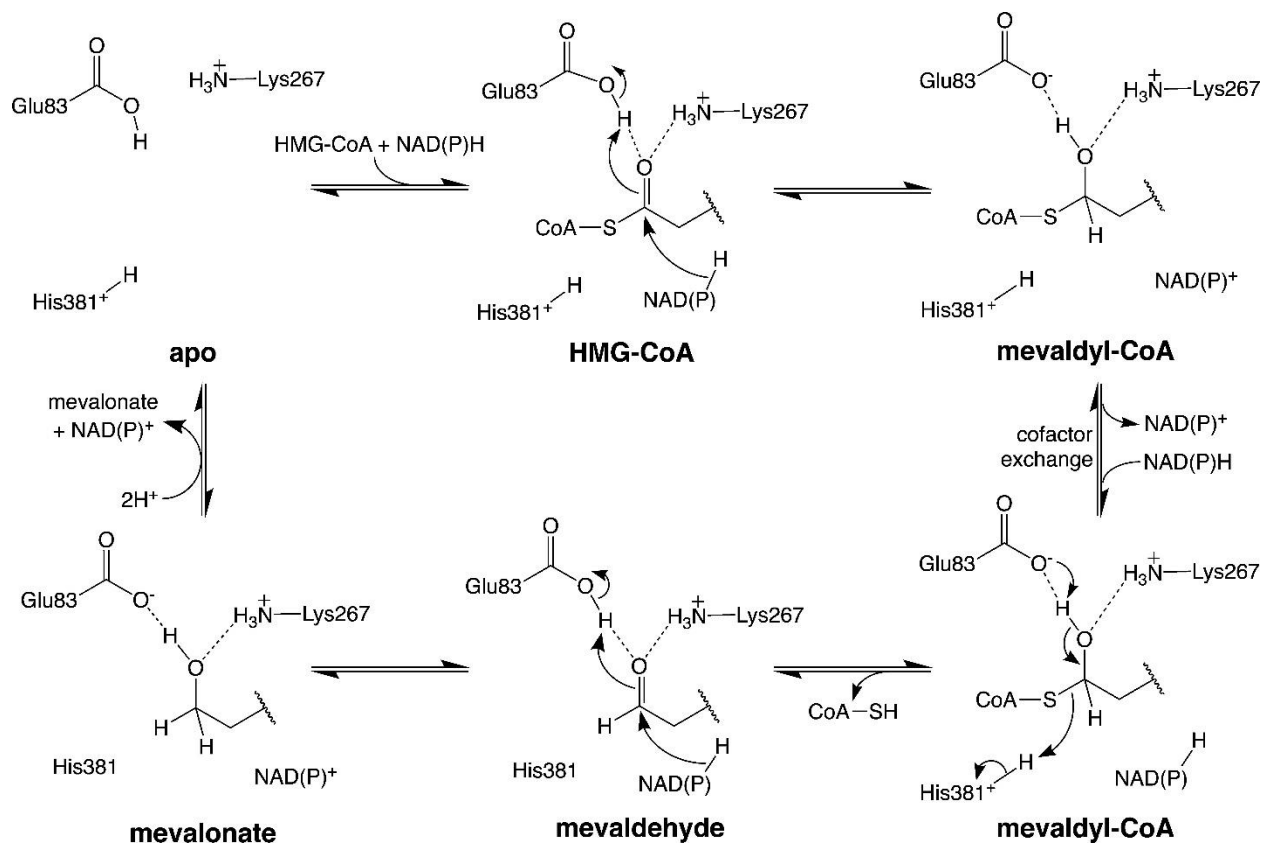


**Figure 18. Two new open conformations of the CTD found in SpHMGR mutants.** Surface representation of the main body of HMGR shown in gray. CTDs are shown in cartoon representation. The open conformation of the CTD as observed in prior SpHMGR structures (PDB 5WPJ, PDB 3QAU), shown in different shades of gray. Yellow = NAD-bound SpHMGR\_DaHMGR-helix, and NAD-bound SpHMGR\_DaHMGR-helix\_M185T. Substrate and cofactor are modeled in and shown in spheres to help situate the CTD with respect to the active site.

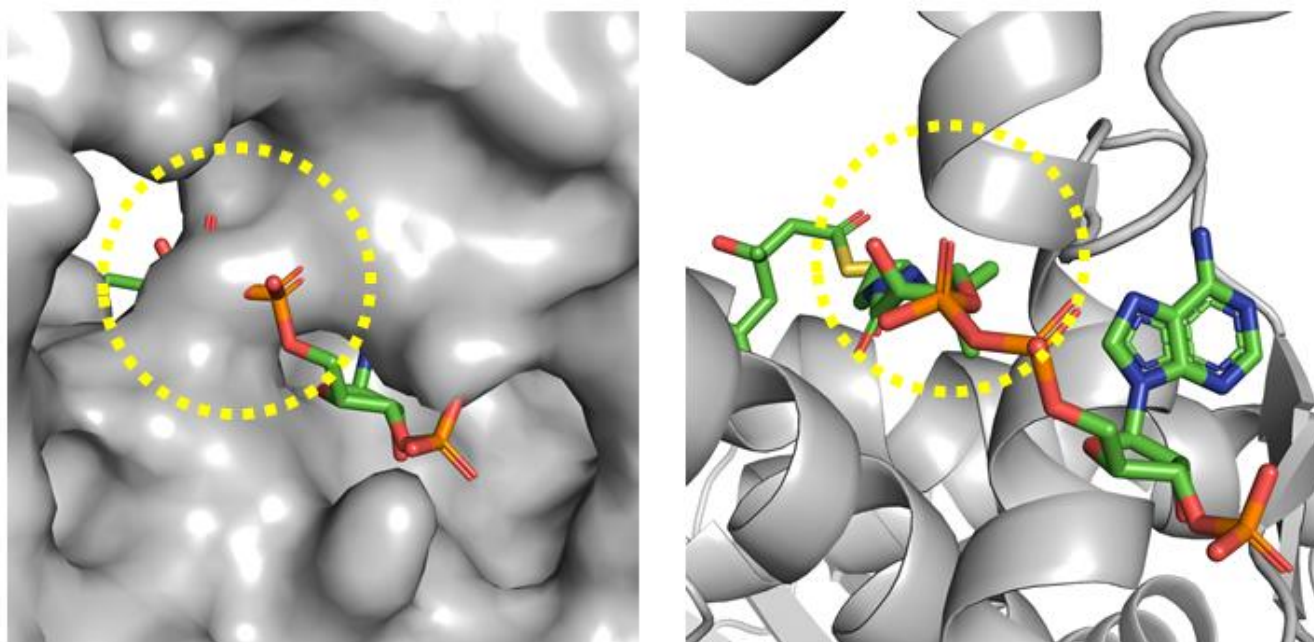


**Figure 19. Flipped conformations of the C-terminal domain as observed in structures obtained in this research.** Surface representation shown in gray. Substrate and cofactor are modeled into the active site and shown as spheres. The CTDs are shown in cartoon representation and are colored as follows: darker blue, darker red, and darker green are the three helices from proceeding from the N to C termini of NADP- and citrate-bound SpHMGR and lighter blue, lighter red, and lighter green represent the analogous helices from NAD- and citrate-bound DaHMGR.



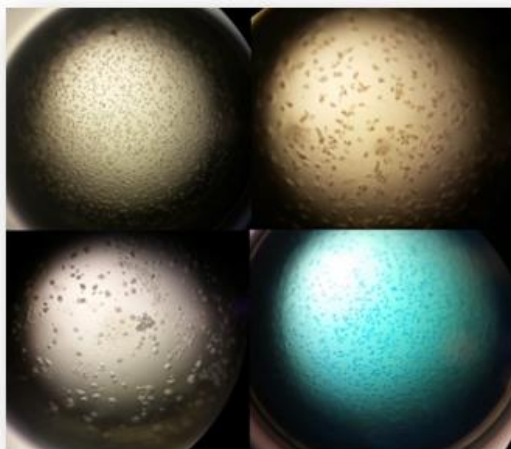


**Figure 20. Revised proposed mechanism.** class II HMGR mechanism, as proposed by Ragwan, E., Arai, E., Kung, Y., *Biochemistry*. 2018

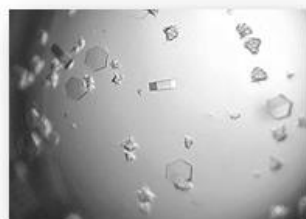


**Figure 21. Steric clash between CoA moiety and the C-terminal domain in the flipped conformation in DaHMGR.** Surface representation (left) and cartoon representation (right) of DaHMGR with the CTD in the flipped conformation. CoA, modeled in from HMG-CoA-bound PmHMGR (PDB 1QAX) and shown as sticks. The steric clash between the CTD and the CoA moiety is highlighted in yellow.

**irregular crystal morphologies in various crystallization conditions**



**optimized citrate condition,  
yielding three-dimensional  
hexagonal crystals**



**optimized Bis-tris condition  
resembling hexagonal crystals that  
are halved**



**Figure 22. Optimization of various DaHMGR crystals.** In the panel on top, there are images of various irregular crystals formed in many different crystallization conditions after initial high-throughput sparse matrix screening. The bottom right of the panel on top, is an image of the crystal drop being dyed with IZIT dye showing that these irregular crystals absorbed the dye which would suggest that they were protein crystals. However, most of these could not be initially optimized and were not ideal for X-ray diffraction experiment. A citrate condition was optimized, as shown in the middle above, revealed a hexagonal crystal morphology. These crystals resulted in the first published citrate- and NADH-bound structure (PDB 6DIO). The constituents of this condition were further modulated to reveal a half-hexagonal crystal as depicted in the bottom image. These crystals resulted in the additional structures published of apo-DaHMGR and mevalonate-bound DaHMGR (PDB 6EEU and 6EEV respectively).

**Table 1.** Data collection and refinement statistics for structures of DaHMGR that are published in the Protein Database.

	<b>NADH- and citrate-bound DaHMGR</b>	<b>mevalonate-bound DaHMGR</b>
<b>PDB ID</b>	6DIO	6EEV
<b>Data Collection</b>		
<b>Beamline</b>	APS, 24-ID-E	APS, 24-ID-C
<b>Wavelength (Å)</b>	0.9792	0.9791
<b>Space group</b>	<i>P</i> 6522	<i>P</i> 321
<b>Unit cell</b>		
<b><i>a=b, c</i> (Å)</b>	113.00, 437.48	100.37, 75.81
<b>Resolution (Å)</b>	97.86–2.14 (2.18–2.14)	57.12–1.49 (1.52–1.49)
<b>Wilson B (Å<sup>2</sup>)</b>	21.39	14.85
<b>Total reflections</b>	675,186 (34,374)	267,310 (13,124)
<b>Unique reflections</b>	92,208 (4,476)	70,102 (3,542)
<b>Multiplicity</b>	7.3 (7.7)	3.8 (3.7)
<b>Completeness (%)</b>	100.0 (100.00)	97.6 (99.7)
<b>Mean <i>I</i>/<math>\sigma</math>(<i>I</i>)</b>	5.2 (2.4)	10.5 (2.0)
<b><i>R</i>merge</b>	0.287 (0.843)	0.067 (0.629)
<b><i>R</i>meas</b>	0.309 (0.904)	0.078 (0.735)
<b>CC1/2</b>	0.986 (0.40)	0.993 (0.700)
<b>Refinement</b>		
<b><i>R</i>work</b>	0.1905 (0.2945)	0.1351 (0.2131)
<b><i>R</i>free</b>	0.2389 (0.3378)	0.1646 (0.2632)
<b>r.m.s.d. bonds (Å)</b>	0.007	0.005
<b>r.m.s.d. angles (°)</b>	0.905	0.821
<b>Number of atoms</b>	9,429	3,228
<b>Average B factors (Å<sup>2</sup>)</b>		
<b>All atoms</b>	24.23	23.73
<b>Protein</b>	23.27	20.14
<b>NADH/citrate/mevalonate</b>	25.48	41.4
<b>Water</b>	30.11	37.3
<b>Ramachandran analysis</b>		
<b>Favored (%)</b>	97.14	97.3
<b>Allowed (%)</b>	2.86	2.69
<b>Outliers (%)</b>	0	0
<b>MolProbity Clashscore</b>	5.8	1.41

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

**Table 2. Data collection and refinement statistics for SpHMGR and DaHMGR structures that are not yet published in the Protein Database.**

	<b>NADPH- and citrate-bound SpHMGR</b>	<b>citrate-bound DaHMGR</b>
<b>Data Collection</b>		
<b>Wavelength</b>	0.97918	0.97918
<b>Resolution range</b>	57.47 - 1.92 (1.989 - 1.92)	98.41 - 2.45 (2.55 - 2.45)
<b>Space group</b>	<i>P</i> 32 2 1	<i>P</i> 21 3
<b>Unit cell</b>	143.07 143.0796.510990 90 120	139.17 139.17139.1790 90 90
<b>Total reflections</b>	584002 (58392)	330677 (38663)
<b>Unique reflections</b>	86815 (8587)	33462 (3753)
<b>Multiplicity</b>	6.7 (6.8)	9.9 (10.3)
<b>Completeness (%)</b>	99.86 (99.70)	100.0 (100.0)
<b>Mean I/sigma(I)</b>	7.02 (1.23)	7.9 (1.7)
<b>Wilson B-factor</b>	25.48	45.84
<b>R-merge</b>	0.1735 (1.55)	0.186 (1.377)
<b>R-meas</b>	0.1883 (1.683)	0.208 (1.530)
<b>CC1/2</b>	0.993 (0.42)	0.995 (0.469)
<b>Refinement</b>		
<b>R-work</b>	0.1605 (0.2954)	0.1902 (0.2739)
<b>R-free</b>	0.1933 (0.3196)	0.2245 (0.2841)
<b>Number of non-hydrogen atoms</b>	7260	6322
<b>macromolecules</b>	6391	6172
<b>ligands</b>	184	36
<b>solvent</b>	747	124
<b>Protein residues</b>	844	850
<b>RMS(bonds)</b>	0.007	0.002
<b>RMS(angles)</b>	0.91	0.49
<b>Ramachandran favored (%)</b>	98.33	97.03
<b>Ramachandran allowed (%)</b>	1.67	2.97
<b>Ramachandran outliers (%)</b>	0	0
<b>Rotamer outliers (%)</b>	0.75	1.64
<b>Clashscore</b>	5.8	3.22
<b>Average B-factor</b>	31.68	52.61
<b>macromolecules</b>	30.86	52.82
<b>ligands</b>	24.74	43.21
<b>solvent</b>	39.84	44
<b>Statistics for the highest-resolution shell are shown in parentheses.</b>		

**Table 3. DaHMGR enzyme kinetic reaction mixtures.** Enzyme reaction compositions for NADH, NADPH and HMG-CoA steady-state kinetics.

	<b>NADH</b>	<b>NADPH</b>	<b>HMG-CoA</b>
<b>Tris pH 7.4</b>	50 mM	50 mM	50 mM
<b>NaCl</b>	50 mM	50 mM	50 mM
<b>NADH</b>	5-400 $\mu$ M	-	400 $\mu$ M
<b>NADPH</b>	-	10-500 $\mu$ M	-
<b>Enzyme</b>	20 nM	100 nM	20 nM
<b>HMG-CoA</b>	300 $\mu$ M	300 $\mu$ M	2-300 $\mu$ M

**Table 4. HMGR  $K_M$  values of substrates and products observed in various class II HMGRs.** All values are in  $\mu\text{M}$ .

HMGR	HMG-CoA	R,S-Mevalonate	CoA	NAD+	NADH	NADP+	NADPH	Reference
<b>DaHMGR</b>	$7.34 \pm 0.71$				$28.5 \pm 1.4$			
<b>PmHMGR</b>	20	260	60	210	80	52000		[12]
<b>AfHMGR</b>	175	620	30	500	160	1700	500	[21], [22]
<b>LmHMGR</b>	$19.8 \pm 1.2$	$427 \pm 92$	$145 \pm 19$	$25000 \pm 2180$	$150 \pm 11$	$445 \pm 89$	$12.9 \pm 0.9$	[22], [23]
<b>SaHMGR</b>	40	670	390		100	580	70	[22]
<b>SpHMGR</b>	$75.86 \pm 4.43$				153		28.9	[22], [24]
<b>EfHMGR</b>	20	1000	230			250	30	[3], [25]

**Table 5. Summary of the steady-state kinetics of DaHMGR with substrate and cofactor.**

	<b>HMG-CoA</b>	<b>NADH</b>	<b>NADPH</b>
<b>K<sub>m</sub> (μM)</b>	7.34 ±0.71	28.48 ±1.39	N.D.
<b>k<sub>cat</sub> (s<sup>-1</sup>)</b>	22.29 ±0.48	16.58 ±0.24	N.D.
<b>k<sub>cat</sub>/K<sub>m</sub> (M<sup>-1</sup>/s<sup>-1</sup>)</b>	3.04 x 10 <sup>6</sup>	5.82 x 10 <sup>5</sup>	N.D.



### 3.8. References

- [1] L. Sellés Vidal, C. L. Kelly, P. M. Mordaka, and J. T. Heap, “Review of NAD(P)H-dependent oxidoreductases: Properties, engineering and application,” *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1866, no. 2, pp. 327–347, Feb. 2018, doi: 10.1016/j.bbapap.2017.11.005.
- [2] B. E. Haines, O. Wiest, and C. V. Stauffacher, “The Increasingly Complex Mechanism of HMG-CoA Reductase,” *Acc. Chem. Res.*, vol. 46, no. 11, pp. 2416–2426, Nov. 2013, doi: 10.1021/ar3003267.
- [3] B. R. Miller and Y. Kung, “Structural Features and Domain Movements Controlling Substrate Binding and Cofactor Specificity in Class II HMG-CoA Reductase,” *Biochemistry*, vol. 57, no. 5, pp. 654–662, Feb. 2018, doi: 10.1021/acs.biochem.7b00999.
- [4] N. Qureshi, R. E. Dugan, W. W. Cleland, and J. W. Porter, “Kinetic analysis of the individual reductive steps catalyzed by  $\beta$ -hydroxy- $\beta$ -methylglutaryl-coenzyme A reductase obtained from yeast,” *Biochemistry*, vol. 15, no. 19, pp. 4191–4197, Sep. 1976, doi: 10.1021/bi00664a010.
- [5] D. Veloso, W. W. Cleland, and J. W. Porter, “pH Properties and chemical mechanism of action of 3-hydroxy-3-methylglutaryl coenzyme A reductase,” *Biochemistry*, vol. 20, no. 4, pp. 887–894, Feb. 1981, doi: 10.1021/bi00507a036.
- [6] B. G. Darnay, Y. Wang, and V. W. Rodwell, “Identification of the catalytically important histidine of 3-hydroxy-3-methylglutaryl-coenzyme A reductase,” *Journal of Biological Chemistry*, vol. 267, no. 21, pp. 15064–15070, Jul. 1992, doi: 10.1016/S0021-9258(18)42146-1.
- [7] Y. Wang, B. G. Darnay, and V. W. Rodwell, “Identification of the principal catalytically important acidic residue of 3-hydroxy-3-methylglutaryl coenzyme A reductase,” *Journal of Biological Chemistry*, vol. 265, no. 35, pp. 21634–21641, Dec. 1990, doi: 10.1016/S0021-9258(18)45788-2.
- [8] C. N. Steussy *et al.*, “A Novel Role for Coenzyme A during Hydride Transfer in 3-Hydroxy-3-methylglutaryl-coenzyme A Reductase,” *Biochemistry*, vol. 52, no. 31, pp. 5195–5205, Aug. 2013, doi: 10.1021/bi400335g.
- [9] M. A. Siddiq and V. W. Rodwell, “Bacterial Metabolism of Mevalonic Acid,” *J Bacteriol.*, vol. 93, no. 1, pp. 207–214, Jan. 1967, doi: 10.1128/jb.93.1.207-214.1967.
- [10] L. Taberner, D. A. Bochar, V. W. Rodwell, and C. V. Stauffacher, “Substrate-induced closure of the flap domain in the ternary complex structures provides insights into the mechanism of catalysis by 3-hydroxy-3-methylglutaryl-CoA reductase,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 13, pp. 7167–7171, Jun. 1999, doi: 10.1073/pnas.96.13.7167.

- [11] S. M. Ma *et al.*, “Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases,” *Metabolic Engineering*, vol. 13, no. 5, pp. 588–597, Sep. 2011, doi: 10.1016/j.ymben.2011.07.001.
- [12] E. R. Ragwan, E. Arai, and Y. Kung, “New Crystallographic Snapshots of Large Domain Movements in Bacterial 3-Hydroxy-3-methylglutaryl Coenzyme A Reductase,” *Biochemistry*, vol. 57, no. 39, pp. 5715–5725, Oct. 2018, doi: 10.1021/acs.biochem.8b00869.
- [13] P. D. Adams *et al.*, “PHENIX : a comprehensive Python-based system for macromolecular structure solution,” *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. 2, pp. 213–221, Feb. 2010, doi: 10.1107/S0907444909052925.
- [14] B. E. Haines, C. N. Steussy, C. V. Stauffacher, and O. Wiest, “Molecular Modeling of the Reaction Pathway and Hydride Transfer Reactions of HMG-CoA Reductase,” *Biochemistry*, vol. 51, no. 40, pp. 7983–7995, Oct. 2012, doi: 10.1021/bi3008593.
- [15] *GraphPad Software*. San Diego, California, USA. [Windows]. Available: [www.graphpad.com](http://www.graphpad.com)
- [16] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie, “iMOSFLM : a new graphical interface for diffraction-image processing with MOSFLM,” *Acta Crystallogr D Biol Crystallogr*, vol. 67, no. 4, pp. 271–281, Apr. 2011, doi: 10.1107/S0907444910048675.
- [17] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, “Phaser crystallographic software,” *J Appl Crystallogr*, vol. 40, no. 4, pp. 658–674, Aug. 2007, doi: 10.1107/S0021889807021206.
- [18] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, “Features and development of Coot,” *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. 4, pp. 486–501, Apr. 2010, doi: 10.1107/S0907444910007493.
- [19] N. W. Moriarty, R. W. Grosse-Kunstleve, and P. D. Adams, “electronic Ligand Builder and Optimization Workbench ( eLBOW ): a tool for ligand coordinate and restraint generation,” *Acta Crystallogr D Biol Crystallogr*, vol. 65, no. 10, pp. 1074–1080, Oct. 2009, doi: 10.1107/S0907444909029436.
- [20] F. Madeira *et al.*, “The EMBL-EBI search and sequence analysis tools APIs in 2019,” *Nucleic Acids Research*, vol. 47, no. W1, pp. W636–W641, Jul. 2019, doi: 10.1093/nar/gkz268.
- [21] J. A. Friesen, C. M. Lawrence, C. V. Stauffacher, and V. W. Rodwell, “Structural Determinants of Nucleotide Coenzyme Specificity in the Distinctive Dinucleotide Binding Fold of HMG-CoA Reductase from *Pseudomonas mevalonii* †,” *Biochemistry*, vol. 35, no. 37, pp. 11945–11950, Jan. 1996, doi: 10.1021/bi9609937.

- [22] A. E. Theivagt, E. N. Amanti, N. J. Beresford, L. Taberner, and J. A. Friesen, "Characterization of an HMG-CoA Reductase from *Listeria monocytogenes* That Exhibits Dual Coenzyme Specificity," *Biochemistry*, vol. 45, no. 48, pp. 14397–14406, Dec. 2006, doi: 10.1021/bi0614636.
- [23] D.-Y. Kim, C. V. Stauffacher, and V. W. Rodwell, "Dual coenzyme specificity of *Archaeoglobus fulgidus* HMG-CoA reductase," *Protein Sci.*, vol. 9, no. 6, pp. 1226–1234, 2000, doi: 10.1110/ps.9.6.1226.
- [24] E. I. Wilding *et al.*, "Essentiality, Expression, and Characterization of the Class II 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase of *Staphylococcus aureus*," *J Bacteriol*, vol. 182, no. 18, pp. 5147–5152, Sep. 2000, doi: 10.1128/JB.182.18.5147-5152.2000.
- [25] L. Feng *et al.*, "Specific inhibitions of annonaceous acetogenins on class II 3-hydroxy-3-methylglutaryl coenzyme A reductase from *Streptococcus pneumoniae*," *Bioorganic & Medicinal Chemistry*, vol. 19, no. 11, pp. 3512–3519, Jun. 2011, doi: 10.1016/j.bmc.2011.04.019.
- [26] M. Hedl *et al.*, "*Enterococcus faecalis* Acetoacetyl-Coenzyme A Thiolase/3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase, a Dual-Function Protein of Isopentenyl Diphosphate Biosynthesis," *J Bacteriol*, vol. 184, no. 8, pp. 2116–2122, Apr. 2002, doi: 10.1128/JB.184.8.2116-2122.2002.

## **Chapter IV. Altering the cofactor specificity of class II HMGR using rational design**

### **4.1. Summary**

The prior chapters were primarily geared towards discovery where the goals were to further characterize HMGR using several class II HMGR homologs. A variety of topics that related to the overall structural dynamics of HMGRs as well as the atomic-level detail regarding its mechanism were discussed. One specific topic focused on NAD(H) and NADP(H) cofactor preferences among HMGRs with Chapter Three introducing structural features that govern cofactor specificity. A unique structural motif, termed the cofactor helix, was proposed to be an important contributor towards cofactor specificity. It is observed that specific interactions between the cofactor helix and the cofactor appear to modulate the enzyme's affinity towards either NAD(H) or NADP(H). With the knowledge gained from those chapters, the research now branches into protein engineering, where focus will be geared towards engineering HMGRs with modified cofactor preferences.

In these efforts, we will show successful attempts at altering the cofactor preferences of HMGRs from *Delftia acidovorans* (DaHMGR) and *Streptococcus pneumoniae* (SpHMGR). DaHMGR, which was characterized as an NADH-preferring HMGR in the research presented in Chapter Three, will be modified to now have a cofactor preference for NADPH. This will be achieved by replacing its cofactor helix, which normally contains residues that facilitate NADH binding, with the cofactor helix of SpHMGR. This chimeric protein now prefers NADPH, which will support the proposed work done in the Kung Lab that highlights the importance of the cofactor helix in governing cofactor specificity. Importantly, however, a critical additional mutation will be required to switch cofactor specificity more convincingly from NADH to NADPH in

DaHMGR. This mutation, T189M, in addition to the cofactor helix, switches the cofactor preference of DaHMGR, which will be assessed by kinetic analysis. Similarly, in SpHMGR, its cofactor helix, which normally has specific interactions that facilitate NADPH binding, will be exchanged with the cofactor helix of DaHMGR. This chimeric SpHMGR now has a higher preference for NADH instead of NADPH. Interestingly, the analogous of the T189M mutation which proved to be crucial in enhancing the DaHMGR chimeric protein, will be shown to not have the same impact in SpHMGR specificity in the same manner and its implications will be discussed.

To interrogate the behavior of these chimeric proteins on a molecular level, we successfully obtained crystal structures of the above chimeric proteins. Since these modified proteins were generated using rationale design, we considered it essential to visualize whether the expected mutations manifested in a predictable manner. In both the case of DaHMGR and SpHMGR, along with the chimeric constructs, the crystal structures will provide good evidence of how switching the cofactor helices promotes switched cofactor preferences and will be discussed in detail. Collectively, the research presented here takes an important step forward in equipping us with the ability to modify cofactor preferences among class II HMGRs.

## **4.2. Introduction**

### 4.2.1. NAD(H) and NADP(H)

Nicotinamide Adenosine Dinucleotide (NAD(H)) and its phosphorylated version, nicotinamide adenine dinucleotide phosphate (NADP(H)), are ubiquitous across biology. Whether as a substrate or cofactor, NAD(P)-dependent reactions facilitate processes in a wide scope of cellular functions ranging from redox homeostasis within pathways to immune functions to

processes involved in aging [1]. Of its many essential redox and non-redox functions, NAD(P)(H) has a major role in metabolism owed to its ability to work with enzymes such as oxidoreductases to facilitate the transfer of electrons between molecules [1]. In this capacity, NAD(P)-dependent oxidoreductase enzymes utilize NAD(P)(H) as an electron reservoir since it can accept or donate a hydride.

Owed to their roles related to oxidative stress, aging and cell survival, it is crucial that the redox state, which is governed by the amount of oxidized versus reduced forms of NAD and NADP, be controlled and balanced [2]. The amount of oxidized versus reduced forms of NAD and NADP have a major influence on the types of reactions that they facilitate. These amounts, expressed as ratios, exist due to the observation that NAD and NADP have different physiological roles where low NADH/NAD<sup>+</sup> ratios facilitate oxidative reactions whereas high NADPH/NADP<sup>+</sup> ratios facilitate reductive reactions [2]. The prevalence of a lower NADH/NAD<sup>+</sup> ratios suggest that NAD<sup>+</sup> predominates in the cell and thus serves as a suitable oxidant whereas higher amounts of NADPH over NADP<sup>+</sup> suggests NADPH is the preferred reductant.

Compartmentalization of NAD<sup>+</sup>/NADH and NADP<sup>+</sup>/NADPH redox couples aids in the relative localization of the various forms of these cofactors and is also regulated by, for example, their biosynthetic enzymes[3]. If too little of these cofactors are available, or if there is an imbalance in these ratios, the cellular redox state is altered which impacts metabolic processes that lead to redox stress, energy stress and, consequently, pathologies that contribute towards disease states [3]. Thus, not only are these cofactors important, but their relative abundance, and ratios are crucial to maintaining homeostasis within the cell.

#### 4.2.2. Structural motifs that bind NAD(P)(H)

The binding domain of enzymes that utilize NAD(P)(H) often contain common structural motifs. These motifs enable the enzyme to accommodate the nucleotide binding domain of NAD(P)(H) and other nucleotides such as FAD. One such motif is the Rossmann fold which is made up of a  $\beta\alpha\beta\alpha\beta$  structure that includes the nucleotide binding domain[4]. For example, in many oxidoreductases that have a Rossmann fold, a ~30 residue motif, made up of a  $\beta-\alpha-\beta$  motif, harbors an N-terminal dipole that coordinates the pyrophosphate of NAD(H). Additional structural motifs that bind these nucleotides include the TIM-barrel and dihydroquinoate synthase-like binding folds [5].

A closer look at the Rossmann fold reveals a “fingerprint” region that comprises the nucleotide binding fold with a notable feature made up of a conserved GXGXXG motif, where X represents variable residues, that enables phosphate binding [6]. In addition, NAD(H)-preferring enzymes that contain a Rossmann fold have an important negatively charged residue, aspartate, or glutamate, that informs cofactor binding. These widely conserved negatively charged residues are able to hydrogen bond with the 2'-hydroxyl of the adenine ribose of NAD(H) [7]. This feature helps to distinguish its preference for NAD(H) over NADP(H).

#### 4.2.3. Structural motifs found in class II HMGRs

Friesen et al. were the first to examine the structural features that comprise cofactor binding using the structure of HMGR from *Pseudomonas mevaloni* (PmHMGR)[7]. They noted, early on, that HMGRs do not have a classical Rossmann fold and have a distinctive cofactor binding domain however they point to a few analogous motifs between the cofactor binding domain of HMGR and the classical Rossmann fold. Firstly, they highlight that PmHMGR does not have a GXGXXG motif, a key motif in the fingerprint region found in a Rossmann fold, however Gly186 of a conserved DAMG sequence is analogous to the second glycine of this motif[ 5]. This assertion

came from observing how Gly186, which is located at the N-terminus of a helix whose N-terminal dipole coordinates the negatively charged phosphate moieties of NADH[7]. In addition, as discussed above, in proteins that adopt a Rossmann fold and are specific to NADH, the invariant aspartate or glutamate coordinates the 2-hydroxyl of the adenine ribose[6]. The analogous residue is observed in PmHMGR as Asp146 which explains why PmHMGR prefers NAD(H) over NADP(H) [7]. Figure 1 shows a comparison between the Rossmann fold and the distinctive cofactor binding domain of PmHMGR and notes the analogous motifs discussed above.

Friesen and colleagues further examined the role of Asp146 in governing cofactor specificity. By mutating this aspartate to alanine or glycine, they observed a diminished catalytic efficiency for NAD<sup>+</sup> and an increase catalytic efficiency for NADP<sup>+</sup> [7]. As illustrated in Figure 1, the aspartate directly coordinates the 2-hydroxyl on the adenine ribose, and thus removing this aspartate eliminates this interaction. Moreover, Friesen et al. suggest that with the removal of the negatively charged aspartate, the 2'-phosphate of NADP(H) is better accommodated. In their attempts to further enhance the specificity for NADP<sup>+</sup>, they postulated that by introducing positively charged residues downstream of Asp146, which include residues 147-149 and 192, that these positive residues could interact the negatively charged 2'-phosphate group of NADP(H). Indeed, they noted enhanced catalytic efficiency with NADP<sup>+</sup> when positions Leu148 or Thr192 were mutated to arginine or lysine however they noted that this enhanced specificity was achieved not from an increase in NADP<sup>+</sup> catalytic efficiency but by a decrease in NAD<sup>+</sup> catalytic efficiency [7].

Among class II HMGRs, cofactor preferences differ among various HMGR homologs. While class I HMGRs exclusively utilize NADP(H), in class II enzymes, HMGRs from *Pseudomonas mevalonii* (PmHMGR) [7], *Delftia acidovorans* (DaHMGR) [8], and *Burkholderia*



*cenocepacia* (BcHMGR) [9], all prefer NAD(H) whereas those from *Enterococcus faecalis* (EfHMGR) [10] and *Streptococcus pneumoniae* (SpHMGR) [11] prefer NADP(H). Other HMGRs, such as those from *Archaeoglobus fulgidus* (AfHMGR) [12] and *Listeria monocytogenes* (LmHMGR) [13] appear to have diminished cofactor selectivity owed to their ability to utilize either cofactor comparably. As pointed out above, Friesen et al. introduced some of the key features that modulate cofactor specificity however in their attempts to alter cofactor specificity they were not able to completely switch the preference from NAD<sup>+</sup> to NADP<sup>+</sup> and they concluded that a necessary prerequisite to be able to achieve this required novel structures with NADP<sup>+</sup> bound [7].

In 2018, the Kung Lab published an NADP(H)-bound HMGR from *Streptococcus pneumoniae* (SpHMGR), the first structure of a class II HMGR with NADPH located in the cofactor binding domain. This took an important step forward in being able to compare and contrast the structural features that govern cofactor specificity among class II HMGRs. Historically, among oxidoreductases, understanding the molecular determinants that govern cofactor specificity has largely focused on the region that differentiates NAD(H) and NADP(H), namely the 2-hydroxyl and 2'-phosphate moieties, respectively [5]. Unsurprisingly, when comparing NAD(H) and NADP(H) binding between PmHMGR and SpHMGR, the importance of this region was visualized. Specifically, the Kung Lab introduced the importance of a  $\alpha$ -helix, termed the “cofactor helix” that provides cofactor-specific binding and steric interactions that aid in cofactor selectivity [11].

On the cofactor helix of NADPH-bound SpHMGR, a tyrosine (Tyr144) replaces the aforementioned aspartate and is both bulkier and uncharged for specific reasons. This tyrosine no longer interacts with the 2'-hydroxyl of NAD(H) and, instead, sterically pushes the cofactor to

facilitate the interaction between NADP(H) and downstream residues of the cofactor helix, namely Ser146 and Arg150 [11]. In this configuration, Tyr144 aids in sterically positioning the 2'-phosphate to interact with Ser146 and positions the adenine moiety to undergo pi-pi stacking interactions with Arg150. In contrast, NAD-bound PmHMGR contains nonpolar residues, L148 and L152, that stack up against the adenine moiety.

Along with sequence analysis of various HMGRs, Miller and Kung proposed a model that describes how the cofactor determines cofactor specificity. This pattern, which is relatively consistently observed among class II HMGRs is as follows: in NAD(H)-preferring HMGRs, an aspartate begins the cofactor helix and nonpolar residues create a hydrophobic environment downstream; in NADP(H)-preferring HMGRs, larger uncharged residues such as tyrosine or histidine begins the cofactor helix and charged residues are positioned downstream. The published work presented in Chapter Three served as a first testcase of this hypothesis. By solving the novel structure of an NAD(H)-preferring HMGR, DaHMGR, we could assess whether the above model persisted. Indeed, the interactions between NADH and the cofactor helix of DaHMGR agreed with the above pattern.

Specifically, DaHMGR and PmHMGR, which are NAD(H)-preferring HMGRs, have slightly different cofactor helices, 146-DKVLIGL-152 in DaHMGR and 146 – DQLLNSL-152 in PmHMGR. The critical residues involved in governing cofactor specificity as proposed by the model above involve the first, third and seventh residues of this helix. When comparing these residues, we observe the conserved aspartate and note that positions 148 and 152 are made up of hydrophobic residues. Thus, the overall pattern in the way these HMGR homologs accommodate their preferred cofactor remains. Figure 2 shows a comparison between NAD- and NADP-

preferring HMGRs alongside sequence analysis that shows how the cofactor helix dictates cofactor specificity.

Apart from the cofactor helix mentioned above, there may be additional residues that are involved in enhancing cofactor specificity. The major structural difference between NADP(H) and NAD(H) is the additional phosphate, so it would make sense that the protein would distinguish cofactors at the region that flanks the additional phosphate. Upon close examination, there is indeed an additional pattern that exists in this region. Figure 3 shows the structure of SpHMGR, an NADPH-preferring HMGR, with a methionine residue, Met185, that protrudes towards the NADPH, possibly contributing towards pushing the adenine moiety further towards the cofactor helix and thus facilitating the coordination of this phosphate and Arg150. In the crystal structure, Met185 is depicted as having one conformation, but since methionine has a long flexible sidechain, it is conceivable that this can occupy a greater region than is depicted by the crystal structure. Tyr144, which is part of the cofactor helix, is believed to also be responsible for positioning the cofactor and therefore both Met185 and Tyr144 may contribute towards stabilizing the NADPH phosphate group and orientation of the adenine ring with respect to the cofactor helix. In contrast, DaHMGR, an NAD(H)-preferring HMGR, has a threonine residue, Thr189, in place of Met185 which has a shorter side chain and capable of hydrogen bonding. It is worth noting that this threonine is within hydrogen-bonding distance to the 2'-hydroxyl of NAD(H) in NADH-bound DaHMGR. As a result, the ribose on NADH may swing into and occupy the region in the vicinity of Thr189 and allow the ribose to be coordinated by Asp146 as shown in Figure 3.

When comparing only class II HMGRs, as depicted in the sequence alignment in Figure 3, threonine is conserved across the NADH-preferring homologs, whereas methionine is conserved across NADPH preferring HMGRs. There are other homologs, however, that either have an

alanine or isoleucine present in place of methionine or threonine, and the cofactor preference of some of those homologs present interesting observations. For example, *Archaeoglobus fulgidus* HMGR (AfHMGR) contains an alanine and is considered a dual coenzyme specific HMGR [12]. Additionally, and interestingly, class I HMGRs which are all NADPH-preferring, have a conserved methionine at this location, further reiterating the potential importance of this residue. While class I HMGRs do not have a cofactor helix, which is replaced by a short loop comprised of serine, arginine and phenylalanine that may be responsible for cofactor preference, the possible role of methionine cannot be overlooked. The potential role of the residue at this location has not been previously identified and may prove to be, at least in part, involved in cofactor specificity.

#### 4.2.4. Goals of this section

NAD(P)-dependent oxidoreductases, which utilize NAD(P)(H) as the hydride reservoir have garnered significant interest in metabolic engineering from their ability to utilize unique substrates and synthesize products with desired stereospecific and regiospecificity under ambient conditions [1]. Within these efforts, there is often a need to customize cofactor usage to balance the redox requirements of respective pathways and improve product formation. Integral to customizing cofactor usage is the need to modulate cofactor specificity, giving researchers the ability to respond to redox requirements and ultimately to yield efficient production of desired cofactors. To this end, a greater understanding of the structural determinants that govern cofactor specificity in the cofactor-dependent enzymes can be leveraged in attempts at altering cofactor specificity.

The research in this section follows this course of inquiry whereby the work presented in the prior chapters, along with previously published worked, is leveraged in efforts at customizing the cofactor preference of class II HMGRs. The overall goal is to further the understanding of

cofactor preference among class II HMGRs by using protein engineering to interrogate the motifs that govern specificity. With the knowledge that the cofactor helix dictates cofactor preference, the first aim seeks to switch the cofactor helix between HMGRs, namely DaHMGR and SpHMGR, to switch their cofactor preferences. The proof-of-concept is that by switching the cofactor helices, the respective binding and steric interactions that promote cofactor preference are also switched. A second aim is to explore the threonine and methionine residues in the vicinity of the cofactor helix in NAD(H)- and NADP(H)-dependent HMGRs, respectively, and how they may contribute towards cofactor specificity. The hope is to elucidate whether there are, in addition to the cofactor helix, other residues that enhance the specificity of the non-preferred cofactor. Aims one and two are assessed by kinetic analysis.

The third aim sought to acquire crystal structures of the above chimeric proteins. Since the overall goal is not simply to alter cofactor specificity, but more importantly, to further understand the structural determinants of cofactor specificity, acquiring crystal structures is critical. By acquiring these structures, the hope is to observe whether the mutations mentioned above translate in a predictable manner. For example, do the binding and steric interactions between the cofactor and the cofactor helix in wildtype translate to analogous binding and steric interactions when the cofactor is introduced into a chimeric protein?

As briefly mentioned, the benefits of such interests include being able to alter cofactor selectivity for use in biosynthetic platforms. One such benefit is cost since NAD(H) is cheaper than NADP(H) and is more stable [14]. For example, in one study the NADPH-dependent  $\gamma$ -Diketone Reductase that converts 2,5-hexanedione to (2S,5S)-hexanediol ((2S,5S)) was engineered to improve its NADH utility motivated by the cheaper cofactor costs associated with using NADH cell-free biosynthetic systems [15]. Another benefit is for customizing metabolic pathways to

optimize product formation. In another study, NADH-dependent L-alanine dehydrogenase of *Bacillus subtilis*, was engineered to utilize NADPH and was incorporated as a regenerating enzyme in coupled reactions with NADPH-dependent alcohol dehydrogenases [16]. Yet another benefit is to enable enzymes to effectively use both cofactors instead of one. One such example involved engineering the NAD<sup>+</sup>-dependent *Pseudomonas fluorescens* mannitol 2-dehydrogenase oxidation of polyols to be able to utilize NADP(H) almost as well as NAD(H) [17]. Both the direct applications of being able to manipulate cofactor specificity as well as the value in gaining detailed understanding of the structural features that govern HMGR functionality promise fertile investigation.

### 4.3. Results

#### 4.3.1. Kinetics

##### 4.3.1.1. Kinetic results of DaHMGR, DaHMGR mutants and DaHMGR chimeric constructs

DaHMGR has been shown to significantly prefer NADH over NADPH [8]. This preference is observed through the specific activities in Figure 4 which shows an ~1500-fold higher activity using NADH over NADPH. Upon switching the cofactor helix of DaHMGR to the corresponding helix contained in the NADPH-preferring SpHMGR –a chimera called DaHMGR\_SpHMGR-helix –we observe that NADPH is now slightly preferred over NADH with an ~1.5-fold higher specific activity. An additional mutation further enhances this switch. As previously discussed, many NADP(H)-preferring HMGRs contain a methionine near the phosphate that differentiates NADP(H) from NAD(H) whereas in NAD(H)-preferring HMGRs, smaller residues are observed, as is the case with DaHMGR which contains a threonine (Thr189). After including the additional T189M mutation alongside the SpHMGR-helix – a chimera called DaHMGR\_SpHMGR-

helix\_T189M –there is an even greater preference for NADPH as illustrated by a ~4-fold higher specific activity over NADH.

Since the addition of T189M mutation to DaHMGR\_SpHMGR-helix resulted in a greater switch in cofactor specificity, we questioned whether the T189M alone could switch specificity. To address this, we created DaHMGR\_T189M, DaHMGR with only the single mutation. DaHMGR\_T189M behaves like the wildtype in that it prefers NADH however this mutant shows both a decrease in NADH activity and an increase in NADPH activity which results in a ~680-fold higher specific activity of NADH over NADPH as compared to ~1500-fold higher specific activity of over NADPH in wildtype. This suggests that the T189M mutation alone cannot switch cofactor preference but works in concert with the necessary cofactor helix to influence specificity.

Next, we questioned whether this switched preference, from NADH to NADPH, caused by switching the respective helix, was a result of simply diminishing its ability to use NADH, or whether the switched helix also caused an increase in NADPH activity. To do this, we investigated how the specific activities of the various mutants compared to wildtype using both NADH and NADPH. Figure 5 represents the specific activities of the mutants with respect to wildtype. When assessing activity using NADH, we observe a decline in specific activity as we go from DaHMGR\_T189M, to DaHMGR\_SpHMGR-helix, to DaHMGR\_SpHMGR-helix\_T189M, suggesting that these mutations incrementally diminish NADH preference. When assessing activity using NADPH, we observe an increase in specificity activity as we go from DaHMGR\_T189M, to DaHMGR\_SpHMGR-helix, to DaHMGR\_SpHMGR-helix\_T189M suggesting that these mutations incrementally enhance NADPH preference. Notably, we observe higher specific activity of ~7.6- and ~19-fold respectively in NADPH activities compared to wildtype suggesting that the switched helices markedly improved NADPH usage. Collectively,

this suggests that these chimeric proteins, a switch that relies on switching the helix and is enhanced by the T189M mutation, achieve their modified cofactor preference by not only diminishing their affinity to NADH but also by gaining greater affinity to NADPH.

#### 4.3.1.2. Kinetic results of SpHMGR mutants

SpHMGR has been shown to prefer NADPH over NADH [11]. This preference is observed through the specific activities in Figure 6 which shows ~35-fold higher activity using NADPH over NADH. Upon switching the cofactor helix of SpHMGR to the corresponding helix contained in the NADH-preferring DaHMGR –a chimera called SpHMGR\_DaHMGR-helix –we observed that NADH is now preferred over NADPH with a ~28-fold specific activity.

Like the DaHMGR chimera kinetics, we questioned whether the aforementioned methionine in proximity of the phosphate in NADPH-preferring HMGRs can be used to influence cofactor specificity. In the case of the SpHMGR chimera, which has switched preference from NADPH to NADH, we subsequently mutated methionine (M185) to threonine, the analogous residue found in NADH-preferring DaHMGR, a chimera called SpHMGR\_DaHMGR\_M185T. Unexpectedly, this mutation did not enhance cofactor switch from NADPH to NADH, showing ~9-fold higher specific activity as opposed to ~28-fold as seen in SpHMGR\_DaHMGR-helix using NADH.

In the same way that was done in the DaHMGR chimera kinetics, we questioned whether this switched preference, from NADPH to NADH, caused by switching the respective helix, was a result of simply diminishing its ability to use NADPH, or whether the switched helix also caused an increase in NADPH activity. To do this, we investigated how the specific activities of the various mutants compared to wildtype using both NADH and NADPH. Figure 7 represents the



specific activities of the mutants with respect to wildtype. We observe lower specific activity of ~15 and ~29-fold respectively in NADH activities compared to wildtype suggesting that the switch did not markedly improved NADH usage. However, NADPH specific activity was lower by ~14,000 and ~10,000 respectively, using NADPH. This suggests that while switching the cofactor helix, which resulted in a switched cofactor preference from NADPH to NADH, lowered both NADH and NADPH specific activities, it had a greater diminishing impact on NADPH activity.

#### 4.3.2. X-ray crystal structures

##### 4.3.2.1. DaHMGR mutants

###### 4.3.2.1.1. DaHMGR\_SpHMGR-helix\_T189M+NADPH

This chimera crystallized with one monomer in the asymmetric unit such that its catalytically related adjacent monomer can be generated by crystallographic symmetry. Only density for the residues 3-376, which excludes the C-terminal domain (CTD), was observed, and was modeled with the protein chain. The residues pertaining to the cofactor helix (R145 to L152) and T189 were deleted from the protein chain, and the model was refined using phenix.refine [18], which resulted in positive density that resembled the cofactor helix of SpHMGR and methionine at position 189. Polder  $mF_o - DF_c$  omit density maps with respect to the cofactor helix, T189M, and NADPH were calculated in Phenix and are shown in Figure 8.

The electron density for NADPH spanning the diphosphate moiety to the nicotinamide ring indicates some degree of disorder, however there is good density in the region that encompasses the adenosine portion, including the important 2'-phosphate that differentiates NADH from NADPH. The interactions between NADPH include Asn216 forming hydrogen bonds with the amide group of the nicotinamide ring and interactions with NADPH as it extends away from the active site with the backbone NH groups of Met185, Gly186 as well as a backbone interaction

between diphosphate moiety and Gly330. At the other end of NADPH, the adenosine makes hydrogen bonds and backbone interactions with Asp183, and pi-pi interactions with Arg152. The 2'-phosphate interacts with Arg152 directly in addition to a water-mediated interaction with the sidechain of Ser148. There are additional solvent interactions between waters and various portions spanning the entire molecule, including the adenosine, diphosphate, and nicotinamide moieties.

#### 4.3.2.1.2. DaHMGR\_T189M+NADH

DaHMGR\_T189M+NADH crystallized like the chimera above. T189 was deleted and the model refined, which resulted in positive density that matched methionine. Positive density for NADH was observed, although the density was not continuous. Electron density for most of NADH was observed, however density for the 2'-hydroxyl was lacking. This appears to be evidence of disorder in this region most likely due to the steric hindrance caused by introducing the T189M mutation. Moreover, the ribose of NADH appears to be shifted (<1 Angstrom) away from T189M. Overall, NADH binds DaHMGR\_T189M in a similar manner retaining the interactions found in NADH-bound DaHMGR. Unsurprisingly, T189M did not position NADH in the “up” position, however it would presumably favor the up position to avoid steric hindrances between the cofactor and T189M. Polder  $mF_o - DF_c$  omit density maps with respect to T189M, and NADH were calculated in Phenix and are shown in Figure 9.

#### 4.3.2.2. SpHMGR mutants

##### 4.3.2.2.1. SpHMGR\_DaHMGR-helix+NADH

This chimera crystallized with one monomer in the asymmetric unit such that its catalytically related adjacent monomer can be generated by crystallographic symmetry. Density for the entire chain excluding the last serine were observed and was modeled with the protein

chain. The residues pertaining to the cofactor helix (A143 to R150) were deleted from the protein chain, and the model was refined using phenix.refine [18], which resulted in positive density that resembled the cofactor helix of DaHMGR. Polder  $mF_o - DF_c$  omit density maps with respect to the cofactor helix and NADH were calculated in Phenix are shown in Figure 10. Density consistent with citrate, which was part of the crystallization buffer, was observed at the active site, as previously observed.

As it pertains to the C-terminal domain (CTD), this domain is positioned in an apparent open conformation. While not in the exact configuration as observed in prior SpHMGR structures, for example PDB 3QAU, the three-helical bundle is positioned away from the rest of the protein and allows for the entry and exit of cofactor and substrate. A more detailed analysis of this observation is outlined in Chapter Three. As it pertains to the cofactor helix, the omit density for the sidechain of Asp144 is poor beyond the alpha carbon and suggests a measure of flexibility. Some positive density nearby indicates that this aspartate may assume multiple conformations which is modeled into the final structure. A water molecule positioned between this aspartate and the 2-hydroxyl of the adenine ribose indicates the possibility of a solvent-dependent interaction between these two moieties. As it pertains to the binding of the cofactor, NADH binds in a similar manner to wildtype binding of NADPH with no major differences in its positioning. The introduction of the nonpolar residues, Val148 and Leu150 appear to be positioned to hinder a phosphate from NADPH from binding.

#### 4.3.2.2.2. SpHMGR\_DaHMGR-helix\_M185T+NADH

This chimera crystallized with one monomer in the asymmetric unit in a manner like the SpHMGR mutant above. Density for the residues 3-425 were observed and was modeled with the protein chain. The residues pertaining to the cofactor helix (A143 to R150) and M185 were deleted

from the protein chain, and the model was refined using phenix.refine [18], which resulted in positive density that resembled the cofactor helix of DaHMGR and the M185T mutation. Polder  $mF_o - DF_c$  omit density maps with respect to the cofactor helix, M185T, and NADH were calculated in Phenix are shown in Figure 11. Density consistent with citrate, which was part of the crystallization buffer, was observed at the active site, as previously observed.

Like the SpHMGR\_DaHMGR-helix+NADH mutant above, the CTD is positioned in an open configuration and is discussed in more detail in Chapter Three. Also similarly observed is a water molecule between Asp146 and the 2'-hydroxyl of the adenine ribose that may aid in a solvent-dependent interaction. Unlike in SpHMGR\_DaHMGR-helix+NADH, density for the sidechain of Asp146 is clear. NADH binds in the upper position, like NADPH-binding in wildtype SpHMGR and in the chimera above.

#### 4.3.2.2.3. SpHMGR\_M185T+NADPH

This mutant crystallized with the homodimer in the asymmetric unit. In both chains, density for residues 4-424 were observed and was modeled with the protein chain. M185 was deleted from the protein chain, and the model was refined using phenix.refine [18], which resulted in positive density that resembled the M185T mutation. Polder  $mF_o - DF_c$  omit density maps with respect to the M185T, and NADPH for one chain were calculated in Phenix are shown in Figure 12. Density consistent with citrate, which was part of the crystallization buffer, was observed at the active site, as previously observed.

The CTD assumes a novel location that more resembles the “flipped” conformation as previously observed in citrate- and NADH-bound DaHMGR. The CTD is the in the vicinity of the closed conformation observed in the ternary complex of PmHMGR (PDB 1QAX), but the orientation of the three helical bundle appears to be flipped. This orientation is discussed more in

detail in Chapter Three. NADPH binds in the upper position, like NADPH-binding in wildtype SpHMGR and in the chimeric proteins above.

#### 4.4. Discussion

NAD(P)-dependent enzymes are utilized throughout nature to facilitate redox reactions that span a wide range of functions. The existence of two forms of these cofactors, NAD(H) and NADP(H), points to different physiological functions whereby NAD<sup>+</sup> primarily serves as an oxidant whereas NADPH as a reductant. The homeostasis of cells greatly depends on the ratio of oxidized and reduced forms of these cofactors which highlights the importance of the relative abundance and distribution of each of these cofactors. Due to the need to differentiate between these two forms of the cofactor, enzymes have adopted various structural features that govern specificity. A Rossmann fold, for example, contains a ~30 residue motif made up a  $\beta$ - $\alpha$ - $\beta$  motif that coordinates the pyrophosphate of the cofactor and in NAD-dependent enzymes a conserved aspartate on this motif makes direct interactions with the 2-hydroxyl that distinguishes NAD(H) from NADP(H).

Within class II HMGRs, prior research focusing on structural features that determine cofactor specificity have revealed residues and motifs that confer specificity. The first comprehensive study was done by Friesen et al. and studied NAD bound by PmHMGR. They noted the importance of Asp146 in PmHMGR that coordinated the 2-hydroxyl of the adenosine ribose of NAD. Mutational studies done on this residue and others around this region highlighted the interactions that aid in PmHMGR's selectivity towards NAD(H) and demonstrated how one could introduce promiscuity between cofactors by imposing mutations that enhancing NADP(H) utility. However, with the lack

of an NADP(H)-bound structure, these efforts were unable to completely switch cofactor specificity. Subsequently, the Kung lab solved crystal structures of NADPH-bound SpHMGR which aided in comparing the cofactor binding motifs between NAD(H)- and NADP(H)-dependent class II HMGRs. A cofactor helix within the cofactor binding site, which flanks the region that differentiates NAD(H) and NADP(H), was proposed as an important feature that confers specificity. Following this discovery, the structure of NADH-bound DaHMGR was solved, which was discussed in Chapter Three, and served to bolster the proposal that the cofactor helix helps to govern specificity.

Now, with the above knowledge of the importance of the cofactor helix, this chapter was geared towards furthering our understanding of the structural features that govern cofactor specificity through protein engineering efforts. Within these efforts, the first aim was to observe if engineering enzymes by switching the cofactor helices between HMGRs, namely NADH-dependent DaHMGR and NADPH-dependent SpHMGR, also switched their cofactor specificity. The second aim was to investigate additional residues involved in specificity. These first two aims were primarily assessed through kinetic analysis. The third and final aim was to acquire crystal structures of these mutants to observe if the mutations that were implemented manifested in the way that they were designed to.

In the first aim, engineered chimeras generated by switching the cofactor helices between DaHMGR and SpHMGR indeed switched their cofactor preferences. This is the first reported study showing success at switching the cofactor specificity among class II HMGRs. In the case of DaHMGR, the wildtype enzyme has a strong preference for NADH and significantly low activity using NADPH, whereas the chimeric construct, DaHMGR\_SpHMGR-helix, now prefers NADPH. Notably, this preference is evidenced by not only a decrease in the specific activity using NADH

but also an increase in the specificity activity using NADPH. This suggests that this chimera, with its newly acquired cofactor helix, does not simply have diminished NADH affinity but also has enhanced NADPH affinity.

In the case of SpHMGR, the wildtype enzyme prefers NADPH, whereas the chimeric construct SpHMGR\_DaHMGR-helix, now prefers NADH. Unlike the analogous chimera above, however, SpHMGR\_DaHMGR-helix achieves its cofactor switch by having a greater diminishing influence on the specific activity using NADPH than NADH. This suggests that this chimera has a low affinity for NADH but an even lower affinity for NADPH, which is the basis for its cofactor selection. This result is not surprising. Cahn et al. point out that typically there is greater success when switching the preferences of oxidoreductase enzymes from NAD(H) to NADP(H) than vice versa and they reason that there is significant potential for enhancing binding that the phosphate provides when attempting to switch from NAD(H) to NADP(H) preference[5]. Importantly, however, switching the cofactor helix in both instances resulted in switching the cofactor preference in the chimeras above, a pioneering effort.

The second aim of this section was to investigate additional residues implicated in cofactor preference, specifically the threonine and methionine residues that are in the vicinity of NADH- and NADPH-dependent cofactors respectively. In NADH-preferring HMGR, DaHMGR, a threonine to methionine mutation was proposed to diminish NADH and enhance NADPH preferences. As observed in Figure 5, this is indeed the case which validates that this region is a hotspot related to cofactor specificity. Importantly, this residue alone was not able to switch cofactor specificity. However, it was observed that the chimera, DaHMGR\_SpHMGR-helix\_T189M, which has both the newly acquired cofactor helix in addition to the T189M, had a greater preference for NADPH than just the chimera without the T189M mutation. Therefore, these

two modifications to wildtype DaHMGR, namely the switch in the cofactor helix and the T189M mutation, works together in its ability to achieve its novel cofactor preference.

The results of the analogous mutation in SpHMGR and SpHMGR chimeras differed. In NADPH-preferring SpHMGR, it was proposed that a methionine to threonine mutation would result in diminished NADPH and enhanced NADH preference. As observed in Figure 7, this was not the case. It appears that this mutation, M185T, diminishes the activity of both cofactors comparably, which does not result in a marked alteration in cofactor specificity. Therefore, it was not surprising to observe that the chimera containing this mutation alongside the newly acquired cofactor helix, SpHMGR\_DaHMGR-helix\_M185T, did not have greater affinity for NADH than without this mutation. This chimera still had a switched preference towards NADH but introducing the M185T minimized this switch instead of enhancing it. This was unexpected, and therefore it was highly desirable to visualize the structural underpinning of this observation. The third aim of this section was to solve crystal structures of the above chimeras to understand how these mutations manifested. Especially considering some of the differences in the way DaHMGR and SpHMGR chimeras behaved, this final aim was of paramount importance.

It is worth pointing out that while the above mutations, namely the switch in cofactor preference because of the switch in cofactor helices, were successful, the overall activity of these modified enzymes were significantly lower than wildtype enzymes when they function in their native capacity. In the process of altering cofactor specificity, overall catalytic activity is compromised. This is not an unexpected result. Cahn et. al, who have done the most comprehensive overview of engineering NAD(H)/NADP(H) cofactor preferences among oxidoreductases, point out that even minor changes to the cofactor binding site can have a dramatic effect on overall activity despite the hydride transfer at the catalytic site being distal from the



region that differentiates NAD(H) and NADP(H) binding[5]. To this end, additional work must be done to recover activity.

Concerning the cofactor switch associated with DaHMGR, kinetic analysis revealed that the cofactor switch, from NADH to NADPH, achieved by the chimeric protein, DaHMGR\_SpHMGR-helix\_T189M, was the most effective compared to the chimera without T189M. Thus, we crystallized the former chimera bound to its preferred cofactor, NADPH. The crystal structure of this chimera shows how the binding and steric interactions introduced by engineering DaHMGR to have an NADPH-associated cofactor helix aided in its newly acquired selectivity towards NADPH. Notably, NADPH is bound “higher” up in the cofactor binding than the way wildtype DaHMGR is bound to NADH as will be discussed below.

With the NADPH-associated cofactor helix, the tyrosine lacks the ability to interact with the 2'-hydroxyl of NADH, as illustrated in Figure 13, an overlay of NADH bound in wildtype DaHMGR and NADPH bound in the chimera. Additionally, this tyrosine sterically pushes up the cofactor to facilitate the interaction between the ribose phosphate of NADPH and Arg152, the ribose phosphate and Ser148 through a water molecule, as well as pi-pi stacking interactions between the adenine moiety and Arg152. In addition, there is a second hydrogen bonding interaction between Arg152 and the ribose phosphate as well as a solvent-mediated interaction between the ribose hydroxyl group of NADPH and Ser148. As previously noted, the introduction of the interaction between Arg152 and Ser148 and the phosphate has a major influence on binding. This would explain how this chimera achieves its newly acquired preference by not only diminishing the wildtype preference for NADH but also by increasing its affinity for NADPH. These interactions are very similarly observed in the wildtype NADPH-bound SpHMGR. An overlay of NADPH-bound SpHMGR with this DaHMGR chimera, Figure 14, shows how

remarkably conserved the binding and steric interactions in the NADPH-associated cofactor helix are when translated into DaHMGR.

When examining the role of the T189M mutation, the structure of DaHMGR\_T189M+NADH hints at the role of methionine in influencing selectivity. Without the NADPH-dependent helix, the cofactor binds “lower” in the binding pocket much like the wildtype configuration. However, in this orientation, the methionine creates a less favorable spatial arrangement of the ribose of the cofactor as illustrated Figure 15. This is observed in the omit density maps pertaining to NADH binding, where density for the 2-hydroxyl is lacking and the ribose appears to be shifted (<1 Angstrom) away from T189M (Figure 9). Moreover, the removal of the threonine, eliminates the possible hydrogen bond between the threonine side chain and the 2'-hydroxyl of NADH. With the NADPH-associated cofactor helix, however, the cofactor binds “higher” up in the binding pocket, thus leveraging methionine’s steric role in stabilizing the cofactor in this position. These observations agree with the kinetic analysis which shows that the introduction of T189M reduces NADH activity and enhances NADPH specific activity. Therefore, the structures portray the NADPH-associated cofactor helix being crucial to selecting towards NADPH which is further aided by the T189M mutation.

Concerning the cofactor switch associated with SpHMGR, kinetic analysis revealed that the cofactor switch, from NADPH to NADH, achieved by the chimeric protein SpHMGR\_DaHMGR-helix was the most effective compared to the chimera with the M185T mutation. Thus, we crystallized the former bound to its preferred cofactor, NADH. The crystal structure of this chimera shows how the binding and steric interactions introduced by engineering SpHMGR to have an NADH-associated cofactor helix aided in its newly acquired selectivity towards NADH. However, unexpectedly, NADH is still bound higher up in the cofactor binding site which is discussed below.

With an NADH-associated cofactor helix we expect to see the following: the aspartate directly coordinates the 2'-hydroxyl on the adenine ribose of NADH, non-polar residues creating a hydrophobic environment which would be less favorable for the 2'-phosphate of NADPH, and importantly the lack of a tyrosine to sterically push up the cofactor would allow the cofactor to bind lower in the binding pocket. We do, indeed, observe the hydrophobic residues near adenine ribose of NADH which creates a less favorable spatial arrangement for a phosphate. This likely explains how, with the newly acquired NADH-associated cofactor helix, the kinetic results show this chimera to prefer NADH, and is depicted Figure 16, an overlay of NADPH-bound SpHMGR and NADH bound in this SpHMGR chimera. However, because the cofactor binds higher up in the binding pocket, the aforementioned aspartate does not interact directly with the cofactor. The aspartate, with poor electron density for the side chain possibly indicating multiple conformations, is ~6.5Å away from the hydroxyl of NADH. While there is a possibility that there is a water-mediated interaction between these two moieties, this interaction is different than what was observed in wildtype NADH-bound DaHMGR as shown in Figure 17, an overlay of NADH-bound DaHMGR and this SpHMGR chimera. This may explain why the kinetic results show that while this chimera indeed has a switched cofactor preference, it is primarily achieved by having a greater diminishing influence on NADPH binding than NADH.

As explained above, due to the cofactor binding higher up in the binding pocket, the expected aspartate did not directly coordinate the 2'-hydroxyl of NADH. One hypothesis was that the methionine (M185) could be preventing the cofactor from binding lower. The reasoning for this hypothesis came from the observation that in the analogous DaHMGR chimera, methionine favored the binding of the cofactor in the higher position. However, the structure of SpHMGR\_DaHMGR-helix\_M185T+NADH, with both the DaHMGR helix and the methionine

to threonine mutations, shows that NADH still binds higher up in the binding pocket. Figure 18, shows an overlay of these two novel structures with other structures that depict the cofactor in the higher and lower positions. This figure shows how, despite mutating methionine to threonine in NADPH-bound SpHMGR\_M185T, the cofactor still binds in the upper position. This agrees with the kinetic results which shows that the M185T does not improve cofactor selectivity towards NADH. Instead, since the cofactor still binds higher up in the binding pocket, a methionine and not a threonine, in this position, helps to stabilize the cofactor. The structure of SpHMGR\_M185T+NADPH, which only has the M185T mutation, corroborates this observation by showing that even in wildtype SpHMGR, mutating the methionine to threonine still allows the cofactor to bind in the upper position as illustrated by Figure 18, but due to the lack of the larger sidechain of methionine, this mutant has lower kinetic activity.

To summarize the protein engineering related to DaHMGR, the structural and functional manifestations of the respective mutations can be explained relatively easily. With the introduction of the NADPH-associated cofactor helix, tyrosine aids in pushing the cofactor higher up in the binding pocket where the 2'-phosphate of the cofactor interacts with charged residues. Since the cofactor binds higher in the binding pocket, the threonine to methionine mutation aids in stabilizing the cofactor. Moreover, this configuration no longer possesses the native aspartate to coordinate the 2'-hydroxyl of NADH and the hydrophobic residues to deter phosphate, selectively towards NADPH is achieved.

To summarize the protein engineering related to SpHMGR, the structural and functional manifestations of the respective mutations explain how cofactor selectivity is achieved, but also points to greater complexity. Without additional NADP(H)-bound class II HMGRs structures, it is difficult to draw conclusions, however in the case of SpHMGR, it appears that nature has created

the binding pocket to favor cofactor binding in the higher position. This is likely to accommodate phosphate from its preferred cofactor, NADP(H), requiring additional space. During mutagenesis, aspects of the introduced NADH-associated cofactor helix are translated in SpHMGR, specifically the introduction of hydrophobic residues to deter NADPH binding, and kinetic results agree with this. However, the complete influence of the cofactor helix is diminished since the cofactor helix binds higher up, independent of the cofactor helix.

In this chapter, NADH- and NADPH- selective class II HMGRs were engineering to have altered cofactor specificity, which was assessed by kinetic analysis and visualized using crystal structures. To our knowledge, while many studies attempted cofactor switch in oxidoreductase enzymes, crystal structures of the resultant chimeras are lacking. In a comprehensive review of protein engineering efforts aimed at coenzyme specificity in oxidoreductases, the authors maintain that crystal structures that depict the enzyme with its corresponding cofactor is of significant importance in rational cofactor engineering [19]. The value in crystal structures is to not only validate how these mutations manifest, but to understand how they may behave differently than expected, as was the case in the SpHMGR chimeras. The results of this section provide the first successful attempts at switching the cofactor preferences in class II HMGRs and take a significant step forward in understanding the structural determinants of cofactor selectivity.

## **4.5. Future work**

### 4.5.1. Improving overall catalytic activity

As previously mentioned, modifications to the cofactor specificity in oxidoreductases, including the attempts in this research, results in diminishing the overall catalytic activity of the

enzyme. Cahn et al. propose that following such mutagenesis, it is necessary to recover activity by means of incorporating additional mutations. These mutations serve to “restabilize and reactivate” the protein by compensating for the diminishing effects of the mutations introduced above [5]. The authors note that while the approach to use random mutagenesis screening of various residues even distal to the above mutations result in an effective recovery of activity, it requires significant resources. Instead, based on a database of structural information, they have produced a computational approach whereby one can input their amino acid sequence, and receive recommended residues to target. They note that the most effective mutations are prominently featured in the vicinity of the adenine ring[5].

This computational tool, called Cofactor Specificity Reversal Structural Analysis and Library Design (CSR-SALAD), was used to analyze SpHMGR\_DaHMGR-helix, the SpHMGR chimera which now prefers NADH. CSR-SALAD recommended that the following residues be investigated to aid in recovery of activity: medium priority residues including Thr177, Gln178, Glu179, Ala180, Lys325 and low priority residues, Asp144, Asp176 and Glu179. When analyzing DaHMGR\_SpHMGR-helix\_T189M, the following residues were recommended: medium priority residues Arg182, Asp183, Ala184, and Val328, and low priority residues, Lys151, Arg182, Lys333. Future work can be done in this regard, by introducing mutations with the hope of improving the overall activity of these novel enzymes.

#### 4.5.2. Investigating the structural features that facilitate the cofactor binding in the upper position in SpHMGR constructs

In the case of the DaHMGR chimeric constructs, NADH is bound lower in the binding pocket in the wildtype structure. Upon switching the cofactor helix to an NADPH-associated helix, there are unique interactions that aid in pushing the cofactor higher up in the binding pocket, which

in turn helps to facilitate interactions between the 2'-phosphate and residues of the cofactor helix. In the case of the SpHMGR chimeric constructs, NADPH is bound higher in the binding pocket in the wildtype structure. However, upon switching the cofactor helix to an NADH-associated helix, the cofactor still binds higher, and this prevents, for example, a direct interaction between the aspartate of the cofactor helix and the 2'-hydroxyl of the adenine ribose. It remains unclear why the cofactor still prefers to bind in the upper position.

To this end, it would be valuable to investigate the structural features that govern this phenomenon. If one can modify this region such that the cofactor binds lower in the binding pocket, it is probable that the influence of the cofactor helix will be even more pronounced, thus improving the switching of cofactor preference more convincingly. If it is determined that the cofactor helix in the upper configuration is necessary, then the cofactor helix will need to be modified accordingly. For example, one could attempt using glutamate instead of aspartate to help coordinate the 2'-hydroxyl of the adenine ribose. Since NADH binds higher in the binding pocket, the distance between aspartate and the 2-hydroxyl moieties would be greater when aspartate is used versus glutamate.

#### 4.5.3. Translating and optimizing the helices between additional class II HMGRs

In this research the cofactor helices of two HMGRs were switched resulting in switched cofactor preferences that are determined by the cofactor helix. One could also investigate how the helices of DaHMGR and SpHMGR translate into other class II HMGRs. Does introducing an NADH-associated helix into other NADH-preferring HMGRs retain their cofactor preferences for NADH? Does introducing an NADPH-associated helix into other NADH-preferring HMGRs switch their cofactor preferences to NADPH? Which switches result in more convincing outcomes and what are the structural reasons behind this? This research takes the first significant step in

altering cofactor specificity among class II HMGRs, however, as was demonstrated in the cases of DaHMGR and SpHMGR, each switch needs to be interrogated on a case-by-case basis. Despite this observation, the cofactor helix appears to be the most influential structural feature that governs specificity and thus further work in exchanging and optimizing cofactor-related protein engineering should continue to focus on this motif.

## 4.6. Materials and methods

### 4.6.1. Designing DaHMGR and SpHMGR mutants

The codon-optimized *mva* gene for DaHMGR and SpHMGR wildtype constructs were cloned as previously reported in Chapter Two which contains the nucleotide and amino acid sequences. Round-the-horn PCR mutagenesis was used to insert the cofactor helices to generate the switched-helix mutant constructs. Primers encoding for the SpHMGR cofactor helix (142-AYP SIVKR-150) were used to generate the DaHMGR\_SpHMGR-helix gene and primers encoding for the DaHMGR cofactor helix (145-RDKVLIGL-152) were used for SpHMGR\_DaHMGR-helix. Initially, only the last seven residue were considered part of the cofactor helix however initial kinetic studies showed that eight residues were required. Site directed mutagenesis was utilized to insert the methionine and threonine in the DaHMGR\_SpHMGR-helix\_T189M and SpHMGR\_DaHMGR-helix\_T185M genes respectively. The plasmids were transformed into *E. coli* DH10B cells, and the gene sequences were confirmed (Quintara Biosciences) before transformation into BL21(DE3) cells for protein expression.

### 4.6.2. Expression and purification of DaHMGR and SpHMGR mutants



All DaHMGR and SpHMGR constructs were expressed and purified as reported in their respective wildtype conditions found in Chapter Two. Briefly, cells were grown in lysogeny broth supplemented with kanamycin and was grown at 37°C with shaking until OD<sub>600</sub>~0.6-0.8, induced with 0.5 mM isopropyl β- d-1-thiogalactopyranoside, and allowed to express at 16-20°C at 220 RPM for 16-18 hours. Cells were harvested by centrifugation at 5,000 RPM for 12 minutes, flash frozen with liquid nitrogen, and stored at -80°C. Cells for DaHMGR constructs were sonicated in lysis buffer [50 mM Tris (pH 7.8), 400 mM NaCl, 1 mM tris(2-carboxyethyl) phosphine (TCEP), 10% glycerol, and 10 mM imidazole], clarified, and the supernatant as applied to a Ni-NTA column. DaHMGR constructs were eluted with 250 mM Imidazole, and highest concentration fractions were pooled together, and buffer exchanged in storage buffer [50 mM Tris (pH 7.4), 400 mM NaCl, 1 mM TCEP, and 10% glycerol] using a PD-10 desalting column (GE Healthcare). SpHMGR constructs were purified similarly with the following modifications: lysis buffer [50 mM Tris (pH 7.8), 200 mM NaCl, 10% glycerol, and 10 mM imidazole]; storage buffer [50 mM Tris (pH 7.8), 200 mM NaCl, 10% glycerol]. Protein purity was assessed by SDS-PAGE, with varying concentrations as determined by absorbance at A<sub>280</sub> nm using a calculated extinction coefficient, flash-frozen in liquid nitrogen, and stored at -80 °C.

#### 4.6.3. Kinetic Characterization

Enzyme activity was measured by monitoring the rate of NAD(P)H oxidation by absorbance at A<sub>340</sub> nm using a NanoDrop 2000 with an extinction coefficient of 6200 M<sup>-1</sup> cm<sup>-1</sup>. Reaction mixtures contained 50 mM Tris (pH 7.8), 50 mM NaCl, 5 mM DTT, 300 μM HMG-CoA, 200 μM NAD(P)H, and enzyme. Change in absorbance over time of the initial rates per mg of protein was used to determine the specific activity. Assays were performed in triplicate and the values reported are the means ± the standard error of the mean (SEM).

#### 4.6.4. Crystallization of DaHMGR and SpHMGR mutants

The crystallization condition for Da-SpHelix\_T189M+NADPH was optimized from the wildtype condition previously reported (PDB 6EEU) to contain 100 mM Bis-Tris pH 6.0, 11% PEG, 500 mM LiSO<sub>4</sub> and 10 mM NADPH. Crystals, which looked like the initial condition described in Chapter Two, were grown overnight by sitting-drop vapor diffusion, soaked in a cryo solution containing well solution supplemented with a 51% of a cryo solution containing 16% ethylene glycol, 16% glycerol, 18% sucrose, and 4% glucose (called “magic cryo”) and 20 mM NADPH, and flash-cooled using liquid nitrogen. Crystals for DaHMGR\_T189M+NADH were grown similarly with the addition of 20 mM NADH instead of NADPH.

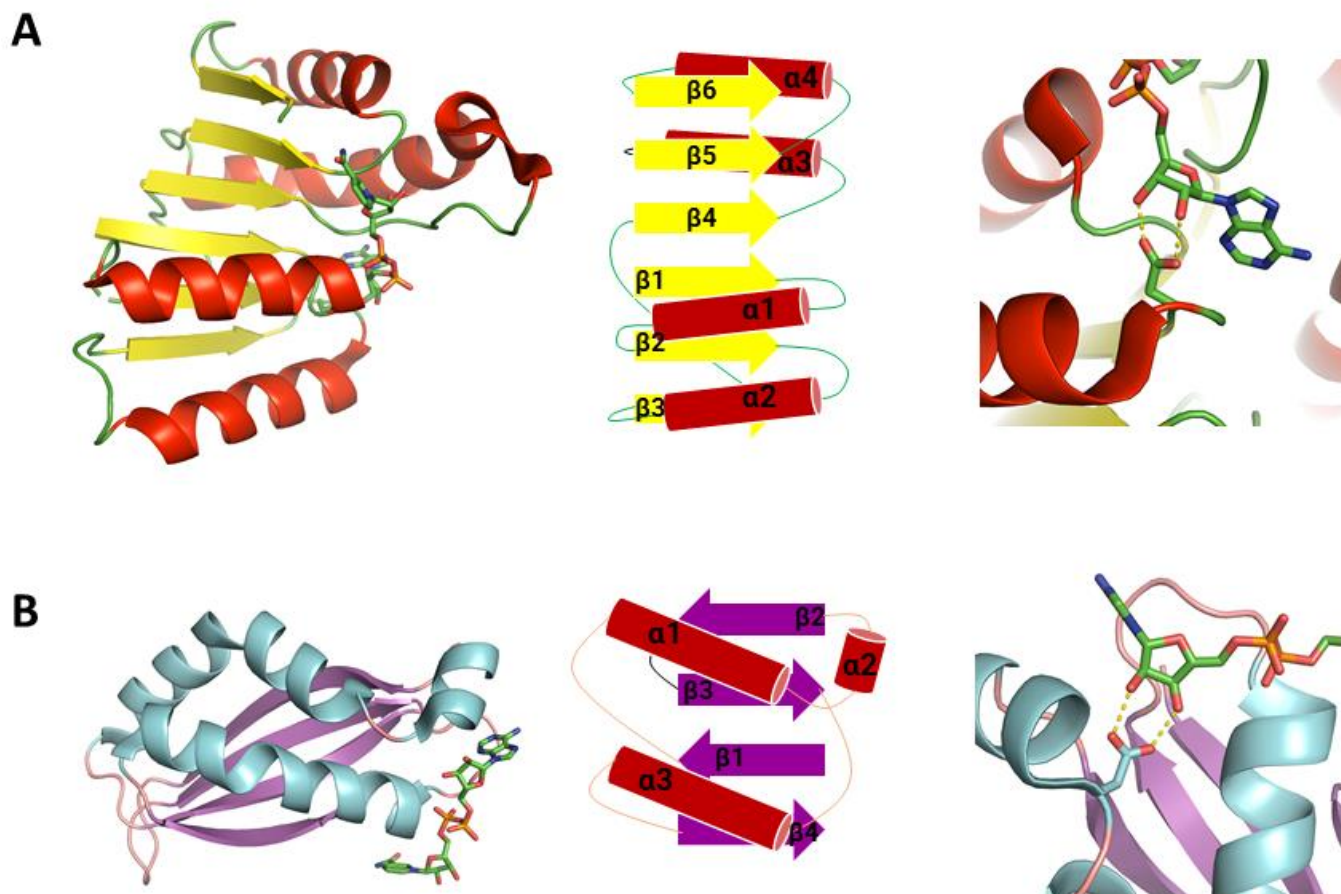
The crystallization condition for SpHMGR\_DaHMGR-helix+NADH was optimized from an initial hit (C5 of PEG<sub>RX</sub> screen from Hampton Research) obtained via sparse-matrix screening by sitting-drop vapor diffusion using a Crystal Gryphon (Art Robbins Instruments) to contain 100 mM Sodium Citrate pH 5.0, 19% PEG 3350 + 20 mM NADH. Crystals, which looked like rectangular sheets, were grown overnight by sitting-drop vapor diffusion, soaked in a cryo solution containing well solution supplemented with 30% glycerol and 20 mM NADH, and flash-cooled using liquid nitrogen. Crystals for SpHMGR\_DaHMGR-helix\_M185T+NADH were grown similarly with the cryo solution contained 35% magic cryo instead of 30% glycerol.

Crystals for SpHMGR\_M185T+NADPH were grown under similar conditions as the SpHMGR constructs above but assumed a different morphology. These crystals, which looked like hexagonal prisms, were grown overnight by sitting-drop vapor diffusion, soaked in a cryo solution containing well solution supplemented with 35% of a solution containing magic cryo and 20 mM NADH.

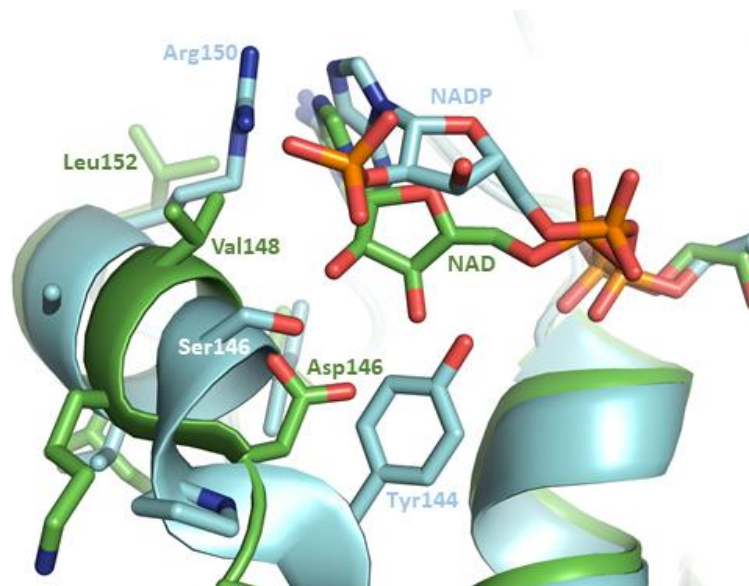
#### 4.6.5. X-ray data collection, structure determination, and refinement of DaHMGR and SpHMGR mutants

X-ray diffraction experiments were conducted at the Advanced Photon Source (APS) beamline 24-ID-E and were indexed, merged, and scaled using iMOSFLM. DaHMGR and SpHMGR structures were solved using apo DaHMGR (PDB 6EEU) and one monomer of a structure of SpHMGR (PDB 5WPJ), respectively, by molecular replacement using Phaser in the Phenix suite. In all cases, ligands and the C-terminal domain were removed from search models. Iterative rounds of model building in Coot and reciprocal refinement in phenix.refine were performed. Where electron density for ligands and CTD were observed, they were modeled in using Coot. Final data collection statistics are listed in Table 1.

## 4.7. Tables and figures

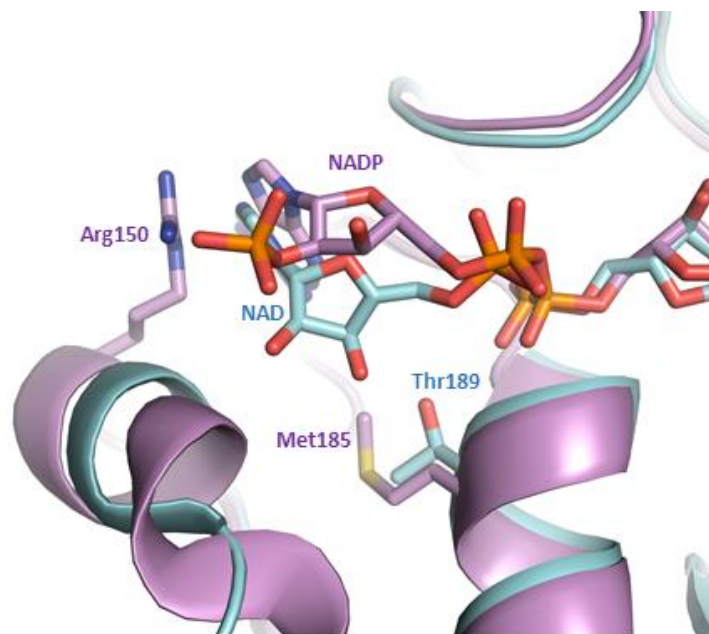


**Figure 1. Comparison between canonical Rossmann nucleotide binding motif versus the nucleotide binding motif of class II HMGR.** On the left are cartoon depictions of the respective nucleotide binding motifs colored by secondary structure. In the middle are schematics of these secondary structures with the  $\alpha$ -helices and  $\beta$ -sheets labeled. On the right are close-ups of the nucleotide binding region showing similarities between the Rossmann fold and the motif found in class II HMGRs that pertain to cofactor selectivity. NAD is shown as sticks and colored as follows: N colored blue, O colored red, P colored orange, C colored green. **A)** Rossmann binding motif as observed in a lactate dehydrogenase structure from *Squalus acanthias* (PDB 1LDM). **B)** The unique binding motif as observed in a class II HMGR structure from *Delftia acidovorans* (PDB 6DIO).



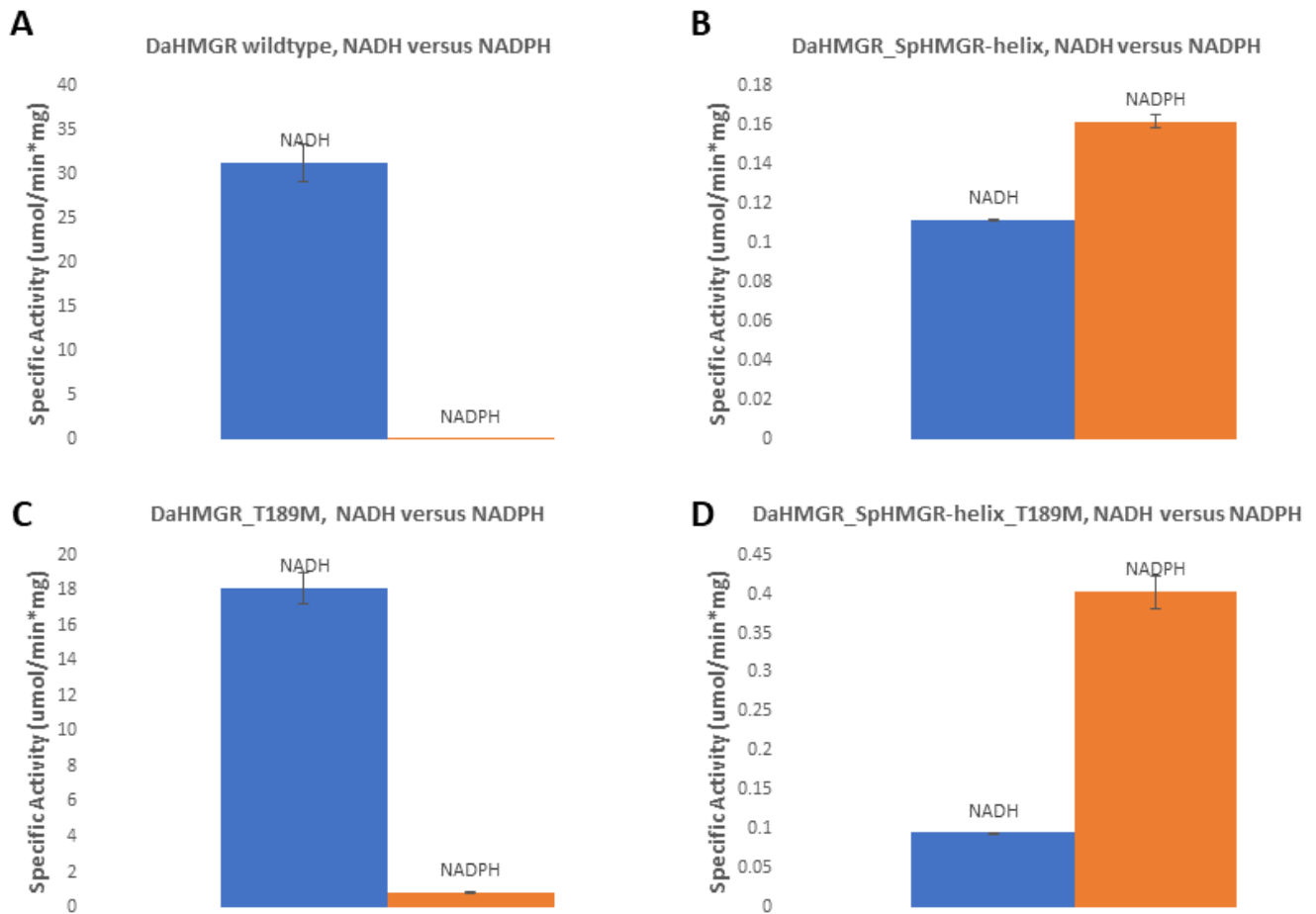
Organism	Cofactor helix	Cofactor preference
<i>Enterococcus faecalis</i>	S Y P S I V K R	NADPH
<i>Streptococcus pneumoniae</i>	A Y P S I V K R	NADPH>NADH
<i>Staphylococcus aureus</i>	A Y P S I K A R	NADPH>NADH
<i>Listeria monocytogene</i>	A H P S L Q K R	NADPH>NADH
<hr/>		
<i>Burkholderia cenocepacia</i>	R D K V L V G L	NADH
<i>Bordetella petrii</i>	C D P I L V K L	NADH
<i>Pseudomonas mevalonii</i>	K D Q L L N S L	NADH
<i>Delftia acidovorans</i>	R D K V L I G L	NADH

**Figure 2. NAD versus NADP associated cofactor helices.** On top, is an overlay of the cofactor helices of NAD-preferring DaHMGR (in green) and NADP-preferring SpHMGR (in cyan) shown as cartoon representation with the cofactors as sticks colored as follows: N colored blue, O colored red, P colored orange, C colored with respect to the color of the cartoon. The first, third, and seventh residues of the cofactor helices are shown in sticks. On the bottom, is a table containing the species, its corresponding cofactor helix, and its cofactor preference. In the table, the first, third and seventh residues of the cofactor helix are highlighted and correspond to the residues depicted as sticks in the diagram above as they are observed in DaHMGR and SpHMGR.

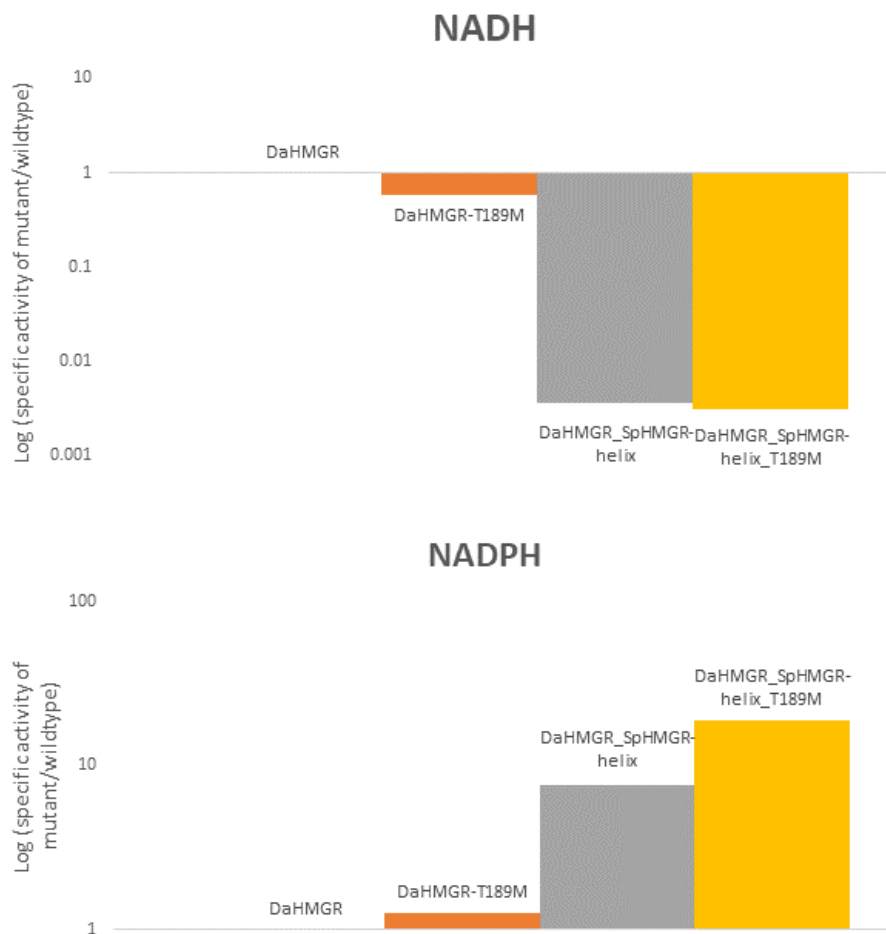


Organism	Sequence	Cofactor preference
<i>Enterococcus faecalis</i>	AMGAN <b>T</b>	NADPH
<i>Streptococcus pneumoniae</i>	AMGAN <b>M</b>	NADPH>NADH
<i>Staphylococcus aureus</i>	AMGAN <b>M</b>	NADPH>NADH
<i>Listeria monocytogene</i>	AMGAN <b>M</b>	NADPH>NADH
<hr/>		
<i>Burkholderia cenocepacia</i>	AMGAN <b>T</b>	NADH
<i>Bordetella petrii</i>	AMGAN <b>A</b>	NADH
<i>Pseudomonas mevalonii</i>	AMGAN <b>T</b>	NADH
<i>Delftia acidovorans</i>	AMGAN <b>T</b>	NADH

**Figure 3. Overlay of the structures of NAD-bound DaHMGR and NADP-bound SpHMGR highlighting residues implicated in cofactor selectivity.** On top, is an overlay of the crystal structures of NAD-bound DaHMGR (in cyan) and of NADP-bound SpHMGR (purple) shown in cartoon representation. The cofactors are shown as sticks as well as the methionine and threonine residues implicated in cofactor selectivity. The sticks are colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring. The methionine appears to push up NADP to facilitate the interaction between the phosphate of NADP and Arg150 (shown as sticks) whereas the overlay suggests that NAD in this position would favor the smaller threonine residue as it is positioned lower in the binding pocket. On the bottom, is a table depicting the organism, cofactor helix and respective cofactor preference of various class II HMGRs showing how larger residues predominate in NADP-preferring HMGRs whereas smaller ones in NAD-preferring HMGRs in the highlighted region that pertain to the methionine and threonine residues depicted in the structures above.

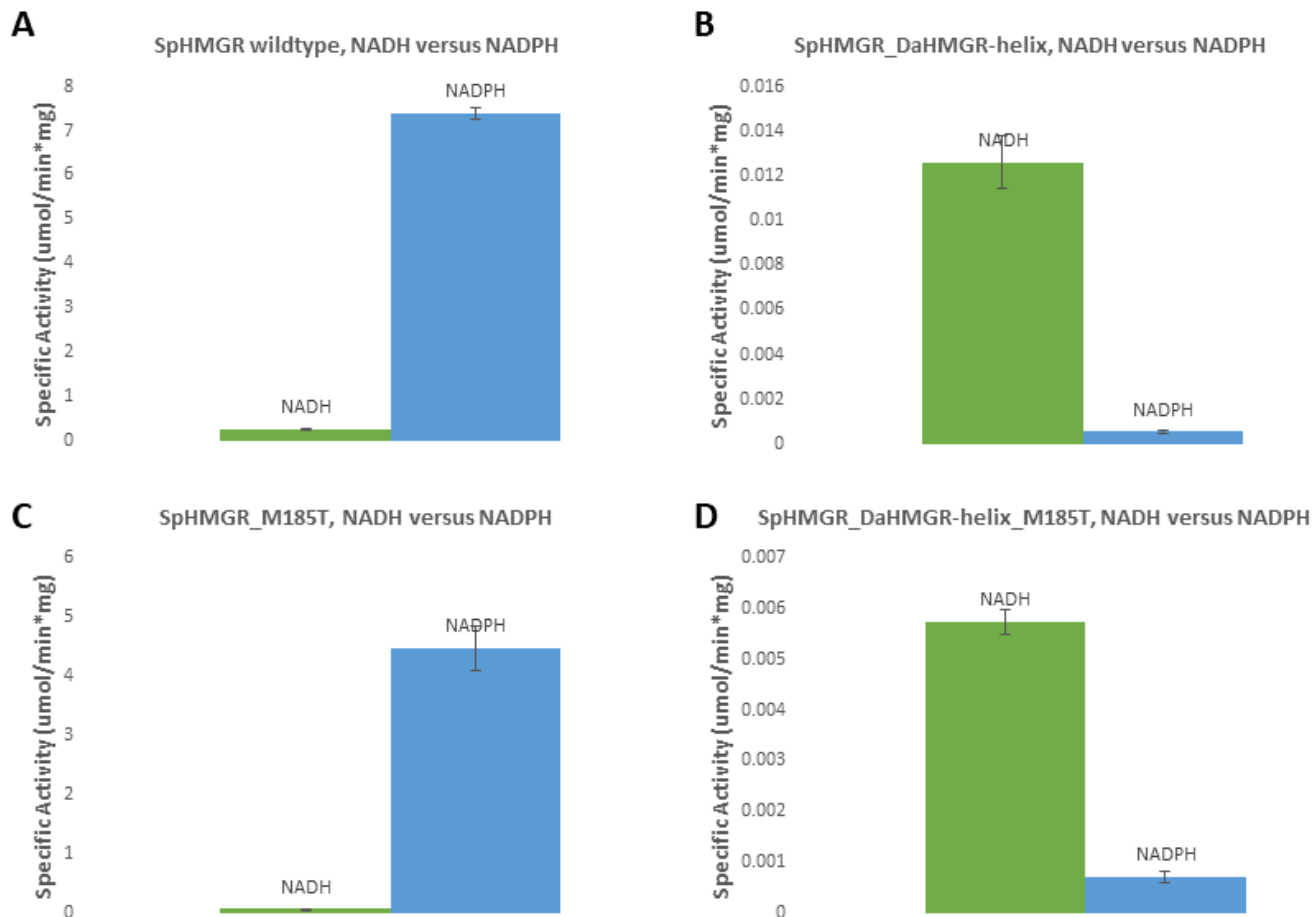


**Figure 4. Kinetics of DaHMGR and DaHMGR mutant constructs.** Specific activity using NADH versus NADPH in **A**) wildtype, **B**) DaHMGR\_SpHMGR-helix, **C**) DaHMGR\_T189M, and **D**) DaHMGR\_SpHMGR-helix\_T189M. Assays were performed in triplicate and the values reported are the means  $\pm$  the standard error of the mean (SEM).

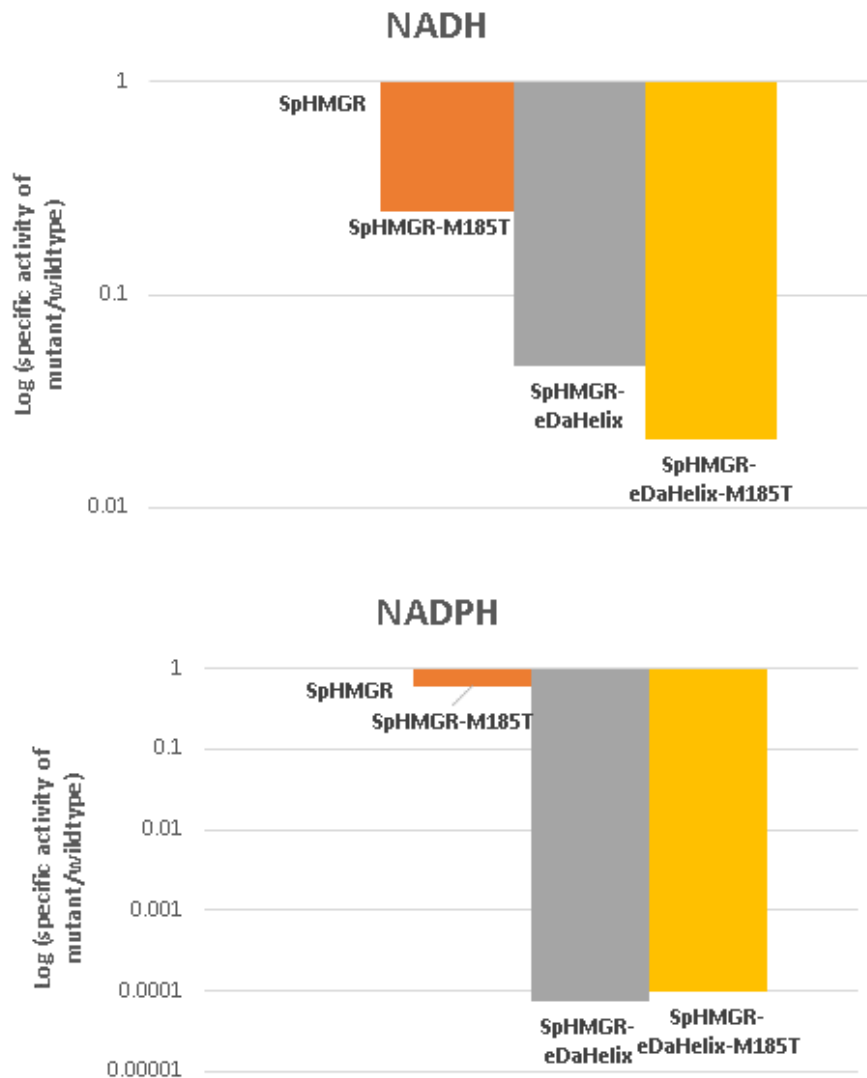


**Figure 5. Comparing activity of DaHMGR mutants versus wildtype with respect to NADH and NADPH.** The specific activities of the various mutants with respect to wildtype DaHMGR are shown on the left, using NADH, and on the right, using NADPH. This was calculated by taking the ratio of the specific activity over the wildtype specific activity and presenting it in a logarithmic scale.





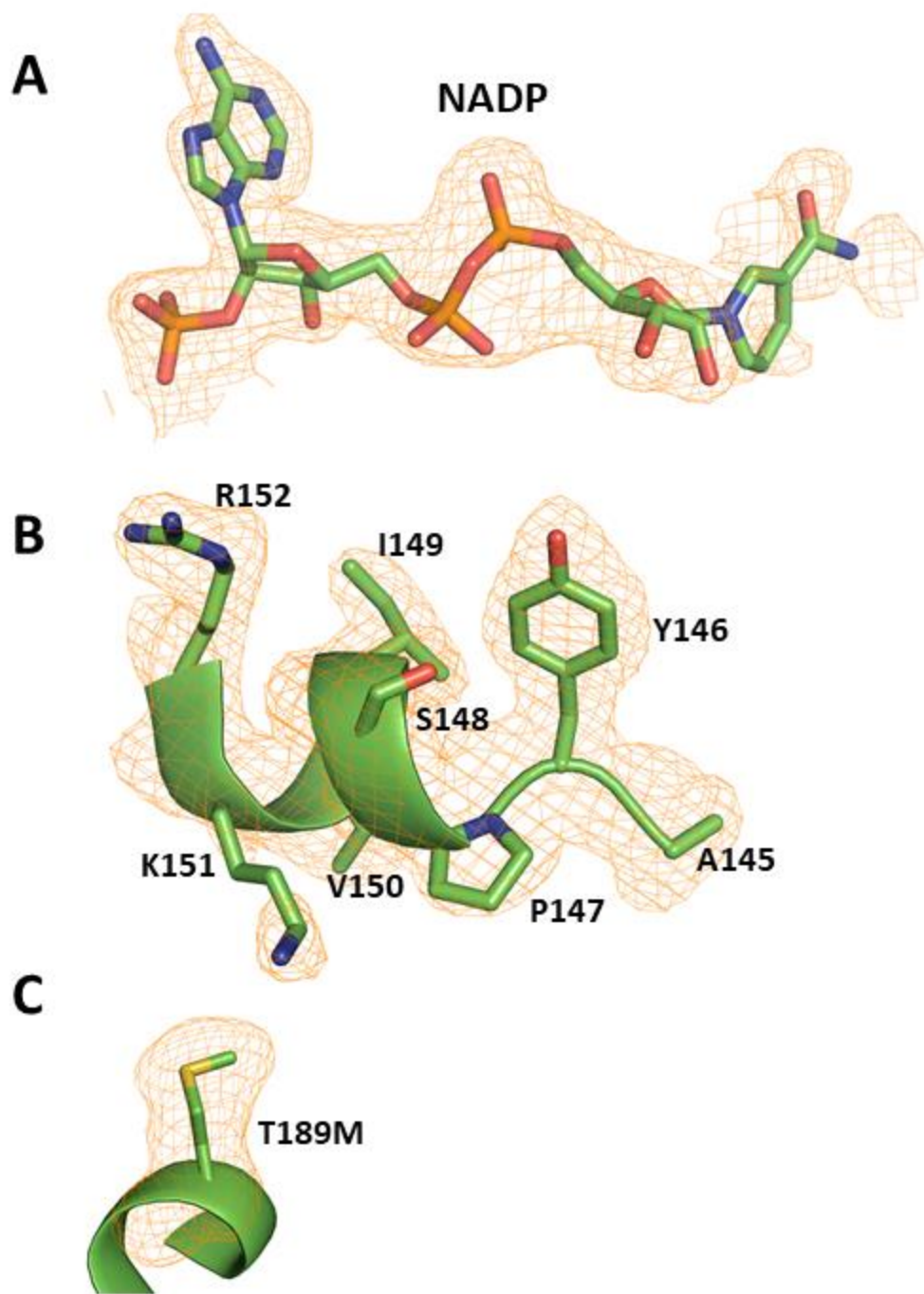
**Figure 6. Kinetics of SpHMGR and SpHMGR mutant constructs.** Specific activity using NADH versus NADPH in **A)** wildtype, **B)** SpHMGR\_DaHMGR-helix, **C)** SpHMGR\_M185T, and **D)** SpHMGR\_DaHMGR-helix\_M185T. Assays were performed in triplicate and the values reported are the means  $\pm$  the standard error of the mean (SEM).



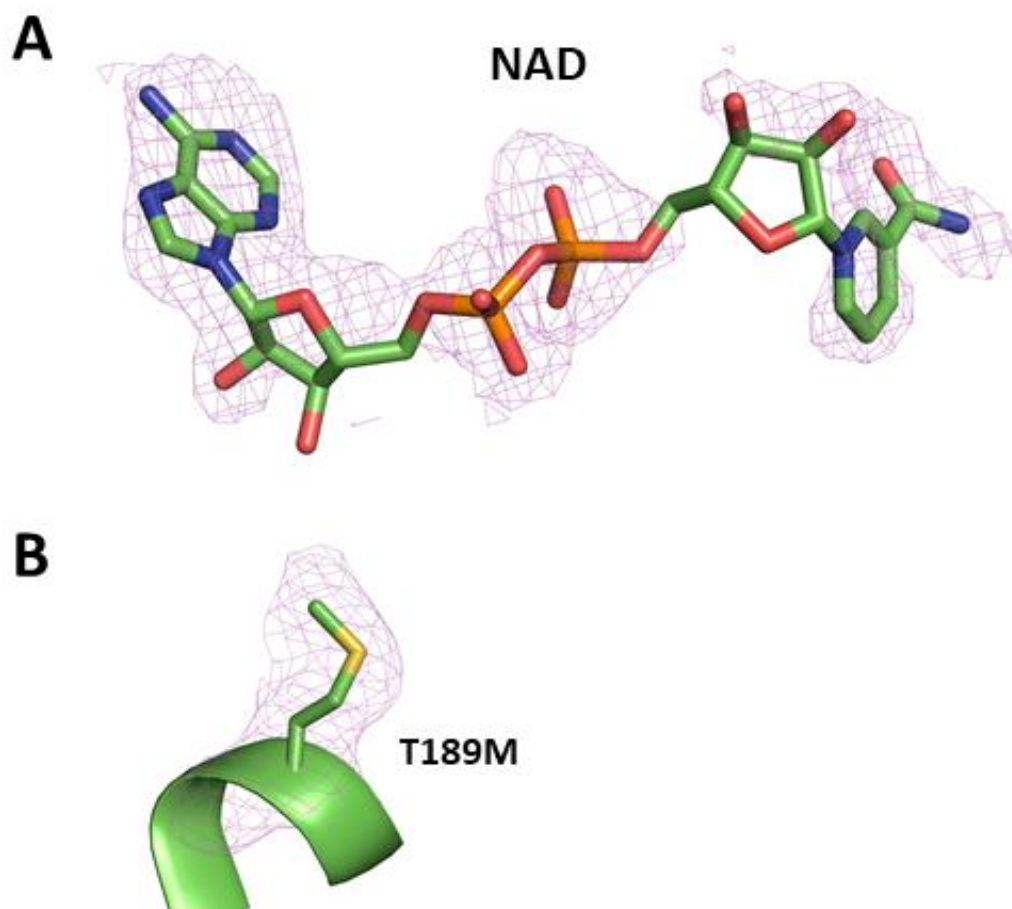
**Figure 7. Comparing activity of SpHMGR mutants versus wildtype with respect to NADH and NADPH.** The specific activities of the various mutants with respect to wildtype SpHMGR are shown on the left, using NADH, and on the right, using NADPH. This was calculated by taking the ratio of the specific activity over the wildtype specific activity and presenting it in a logarithmic scale.

	NADPH-bound DaHMGR_SpHMGR- helix_T189M	NADH-bound DaHMGR_T189M	NADH-bound SpHMGR_DaHMGR-helix	NADH-bound SpHMGR_DaHMGR- helix_M185T	NADPH-bound SpHMGR_M185T
<b>Data Collection</b>					
Wavelength	0.97918	0.97911	0.97918	0.97911	0.97911
Resolution range	75.6 - 1.881 (1.948 - 1.881)	56.94 - 2.24 (2.32 - 2.24)	57.03 - 1.66 (1.719 - 1.66)	53.55 - 1.8 (1.864 - 1.8)	76.13 - 1.53 (1.585 - 1.53)
Space group	<i>P</i> 3 2 1	<i>P</i> 3 2 1	<i>C</i> 2 2 21	<i>C</i> 2 2 21	<i>P</i> 3 2 1
Unit cell	100.22 100.22 75.6 90 90 120	100.21 100.21 75.4504 90 90 120	73.91 89.6501 147.66 90 90 90	73.55 90.29 155.82 90 90 90	141.69 141.69 97.0612 90 90 120
Total reflections	384483 (25342)	105932 (10048)	716151 (35873)	263185 (15494)	1948175 (98616)
Unique reflections	35971 (2298)	20273 (1875)	58349 (2826)	48006 (2827)	168298 (8315)
Multiplicity	10.7 (11.0)	5.2 (5.4)	12.3 (12.7)	5.5 (5.5)	11.6 (11.9)
Completeness (%)	100.0 (100.0)	94.61 (97.43)	99.90 (99.95)	99.15 (99.54)	99.86 (100.00)
Mean I/sigma(I)	11.9 (1.7)	4.5 (1.5)	12.8 (2.5)	6.9 (1.5)	9.5 (1.5)
Wilson B-factor	31.53	26.13	18.62	20.33	15.61
R-merge	0.105 (1.518)	0.234 (0.987)	0.110 (1.052)	0.146 (1.155)	0.151 (1.783)
R-meas	0.116 (1.679)	0.284 (1.204)	0.120 (1.145)	0.178 (1.413)	0.165 (1.952)
R-pim	0.049 (0.712)	0.156 (0.676)	0.047 (0.449)	0.100 (0.798)	0.066 (0.782)
CC1/2	0.999 (0.683)	0.974 (0.276)	0.998 (0.858)	0.993 (0.444)	0.997 (0.527)
<b>Refinement</b>					
Reflections used in refinement	35955 (3558)	20264 (2051)	58087 (5706)	47984 (4751)	168248 (16706)
Reflections used for R-free	1990 (197)	958 (88)	2841 (274)	2385 (228)	8189 (867)
R-work	0.1726 (0.3158)	0.1705 (0.2603)	0.1504 (0.2350)	0.1671 (0.2954)	0.1456 (0.2724)
R-free	0.2063 (0.3493)	0.2116 (0.3017)	0.1742 (0.2723)	0.1943 (0.3418)	0.1787 (0.3036)
Number of non-hydrogen atoms	3024	3010	3536	3576	7304
macromolecules	2751	2732	3216	3195	6425
ligands	108	44	102	57	132
solvent	205	234	257	324	757
Protein residues	375	374	427	423	842
RMS(bonds)	0.011	0.008	0.009	0.007	0.014
RMS(angles)	1.14	1.08	1.04	0.86	1.23
Ramachandran favored (%)	97.32	98.12	98.12	97.62	98.57
Ramachandran allowed (%)	2.68	1.88	1.88	2.38	1.43
Ramachandran outliers (%)	0	0	0	0	0
Rotamer outliers (%)	0.36	0.72	0.91	0.3	0.74
Clashscore	4.09	4.69	4.14	2.63	4.68
Average B-factor	40.36	30.93	29.53	29.82	25.72
macromolecules	39.3	29.9	29.28	28.87	24.7
ligands	71.09	65.15	22.53	33.88	19.71
Solvent	44.3	36.47	34.32	38.39	35.3

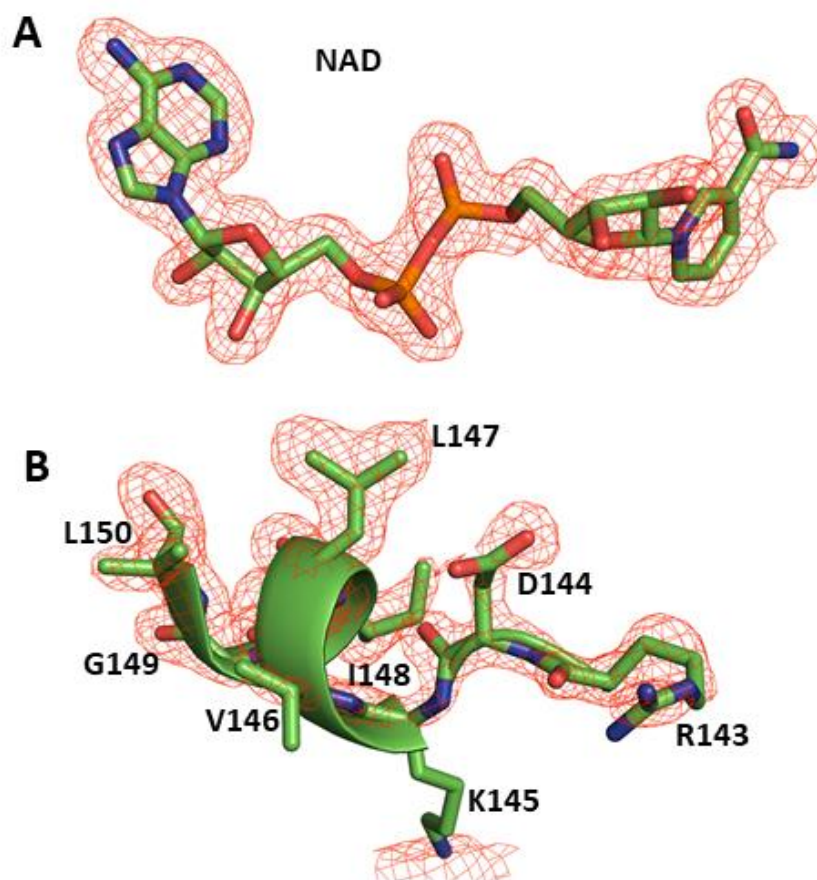
**Table 1. Data collection and refinement statistics pertaining to the various structures in this chapter.**



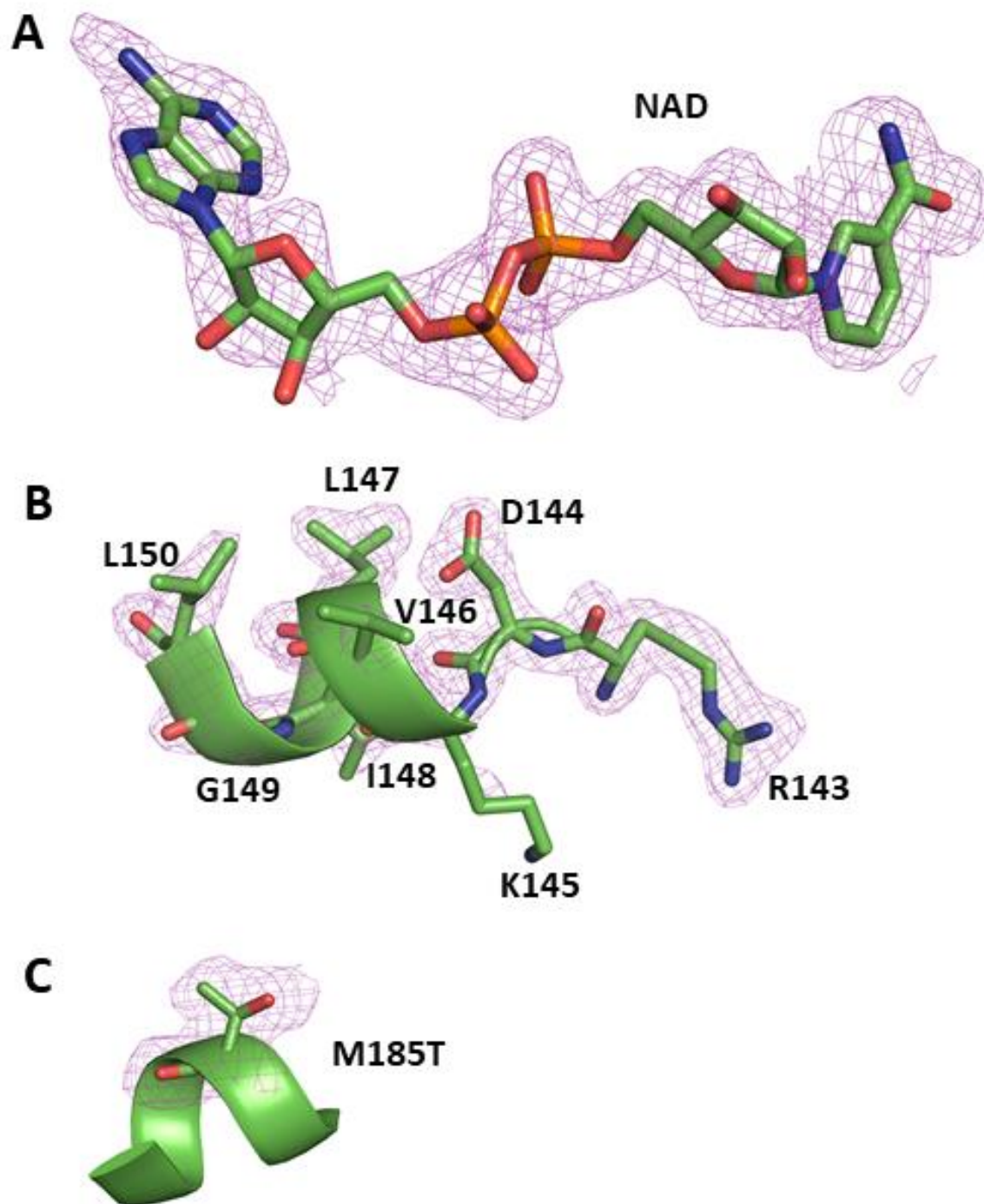
**Figure 8. Polder omit density maps calculated from NADP-bound DaHMGR\_SpHMGR-helix\_T189M.** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  shown surrounding **A)** NADP, **B)** the NADP-associated cofactor helix from SpHMGR, and **C)** the T189M mutation introduced colored orange. Each of the respective moieties are shown in cartoon representation or sticks with the following coloring: N colored blue, O colored red, P colored orange, C colored green.



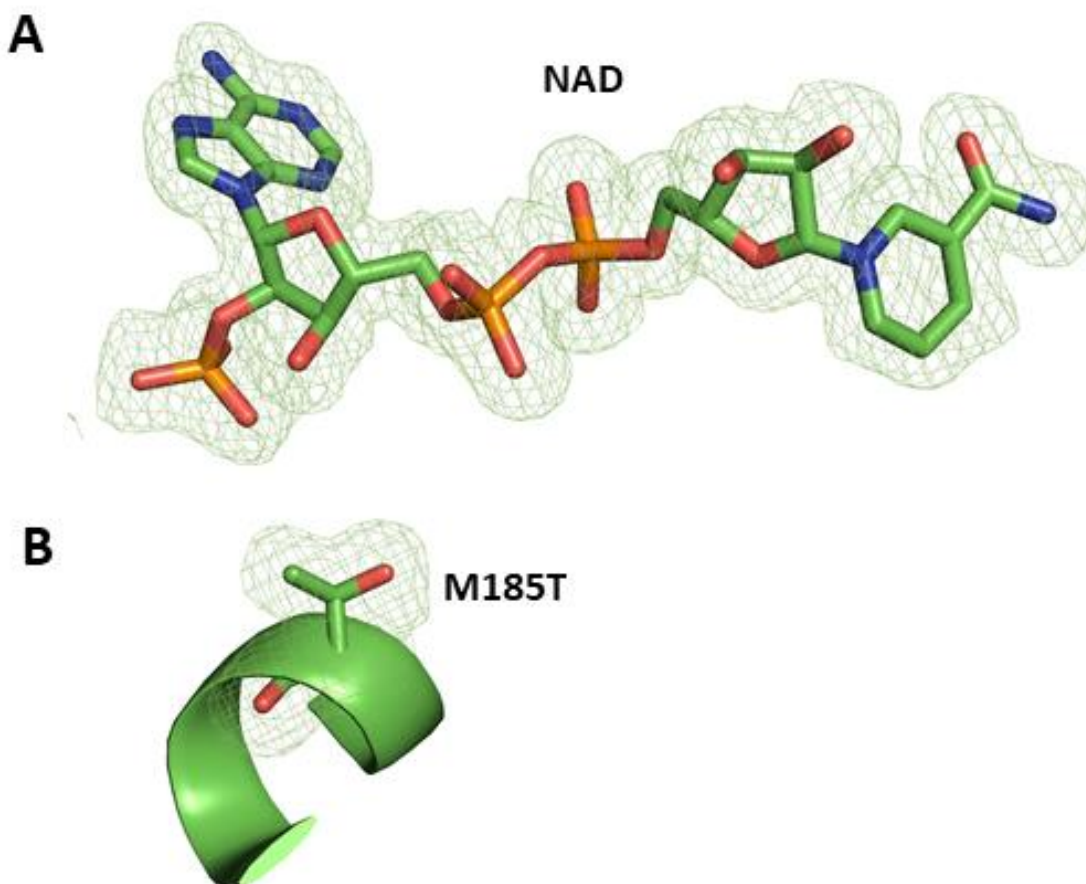
**Figure 9. Polder omit density maps calculated from NAD-bound DaHMGR\_T189M.** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  shown surrounding **A)** NAD, **B)** the T189M mutation introduced colored magenta. Each of the respective moieties are shown in cartoon representation or sticks with the following coloring: N colored blue, O colored red, P colored orange, C colored green.



**Figure 10. Polder omit density maps calculated from NAD-bound SpHMGR\_DaHMGR-helix.** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  shown surrounding **A)** NAD, **B)** the NAD-associated cofactor helix of DaHMGR colored red. Each of the respective moieties are shown in cartoon representation or sticks with the following coloring: N colored blue, O colored red, P colored orange, C colored green.

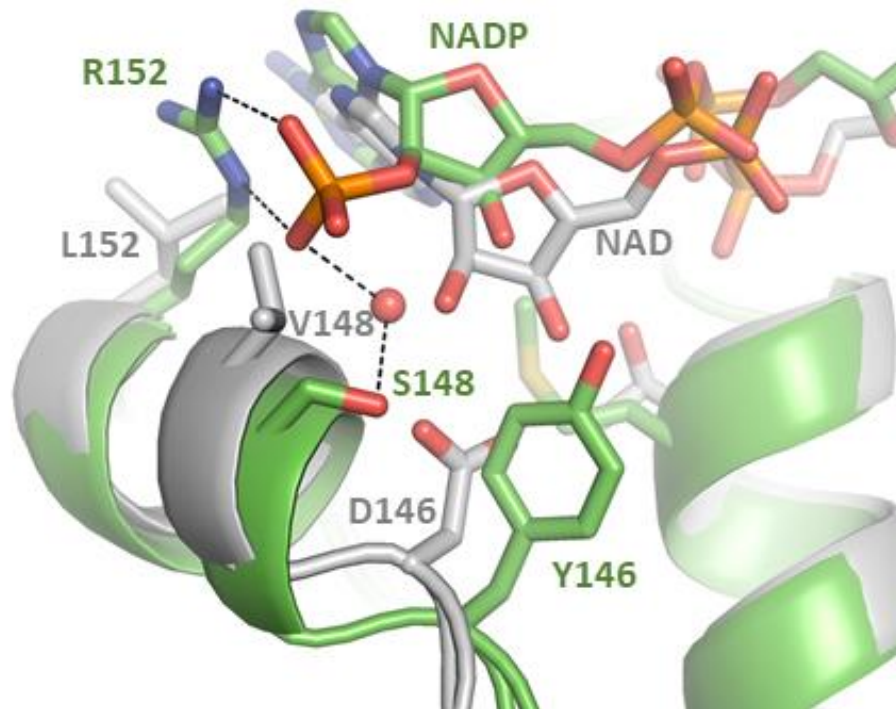


**Figure 11. Polder omit density maps calculated from NAD-bound SpHMGR\_DaHMGR-helix\_M185T.** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  shown surrounding **A)** NAD, **B)** the NAD-associated cofactor helix of DaHMGR, and **C)** the M185T mutation colored purple. Each of the respective moieties are shown in cartoon representation or sticks with the following coloring: N colored blue, O colored red, P colored orange, C colored green.

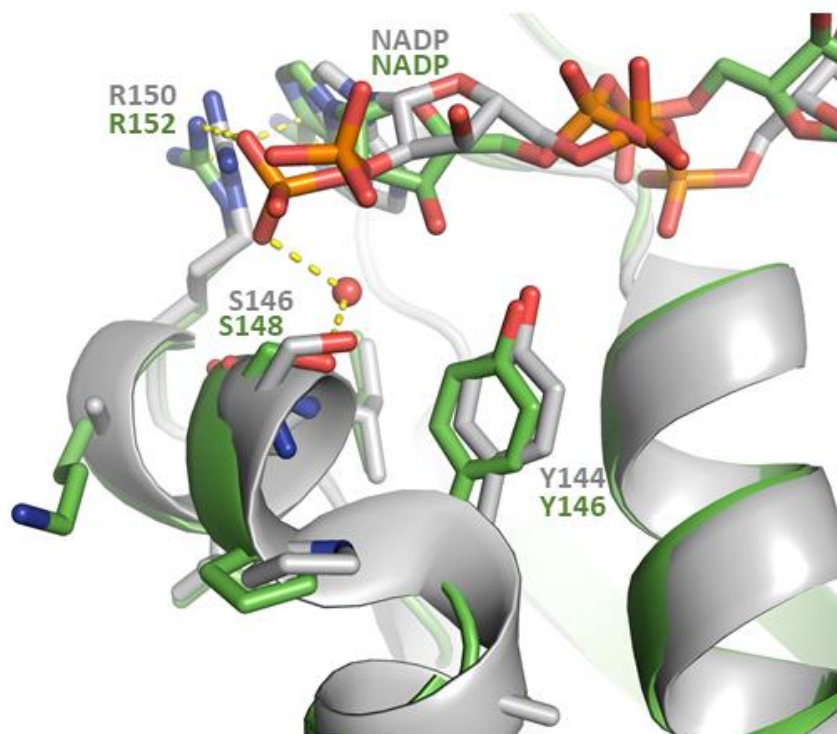


**Figure 12. Polder omit density maps calculated from NADP-bound SpHMGR\_M185T.** Polder  $mF_o - DF_c$  omit density contoured at  $3.0\sigma$  shown surrounding **A)** NADP, **B)** the M185T mutation colored green. Each of the respective moieties are shown in cartoon representation or sticks with the following coloring: N colored blue, O colored red, P colored orange, C colored green.

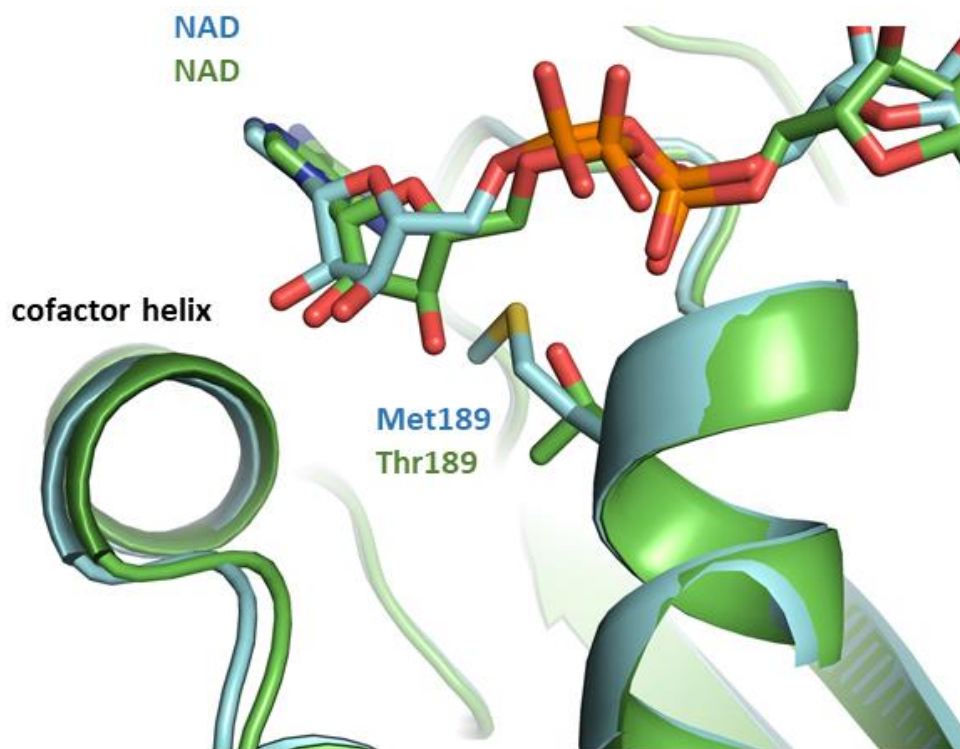




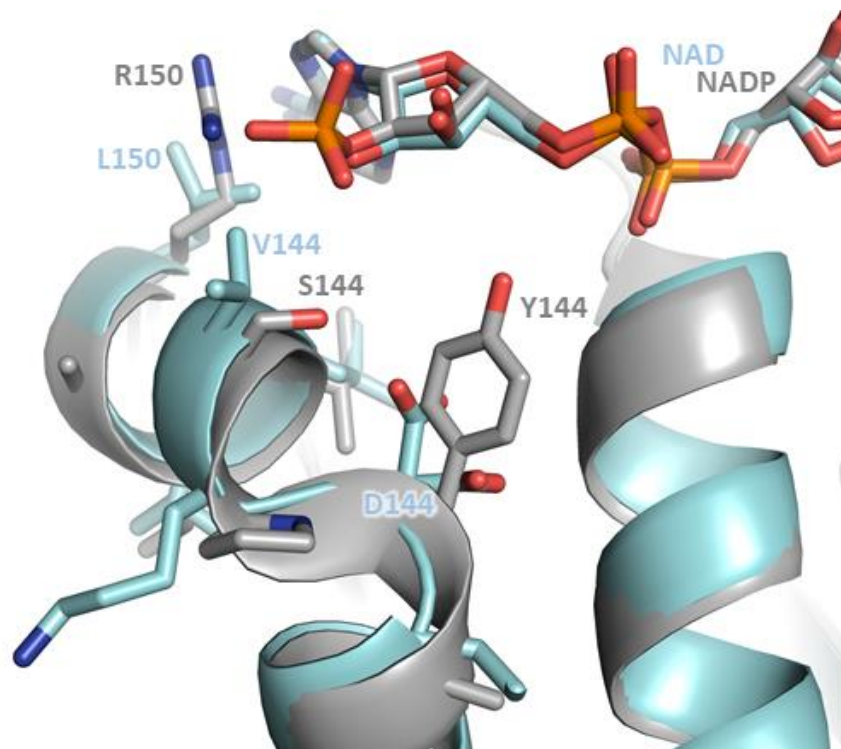
**Figure 13. Overlay of NAD-bound DaHMGR and NADP-bound DaHMGR\_SpHMGR-helix\_T189M.** Overlay of NAD-bound DaHMGR (grey) and NADP-bound DaHMGR\_SpHMGR-helix\_T189M (green) shown in cartoon representation with the cofactor helices and cofactors in sticks colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring.



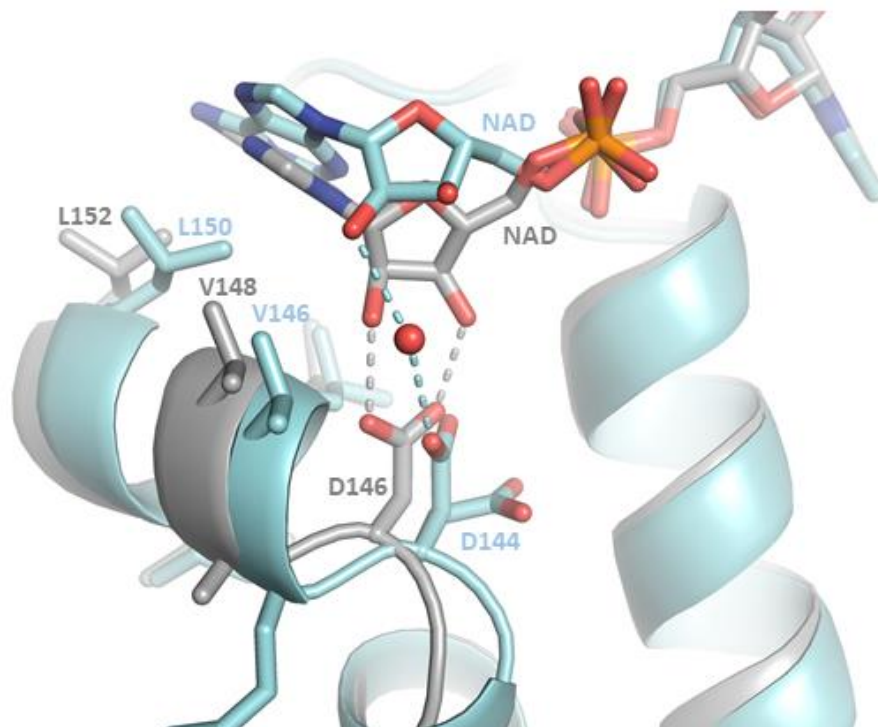
**Figure 14. Overlay of NADP-bound DaHMGR\_SpHMGR-helix\_T189M and NADP-bound SpHMGR.** Overlay of NADP-bound SpHMGR (grey) and NADP-bound DaHMGR\_SpHMGR-helix\_T189M (green) shown in cartoon representation with the cofactor helices and cofactors in sticks colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring. Interactions between Arg152 and the adenine moiety, the 3'-phosphate and Arg152, and the 3'-phosphate and Ser148 through a water molecule are depicted using yellow dashed lines in NADP-bound DaHMGR\_SpHMGR-helix\_T189M.



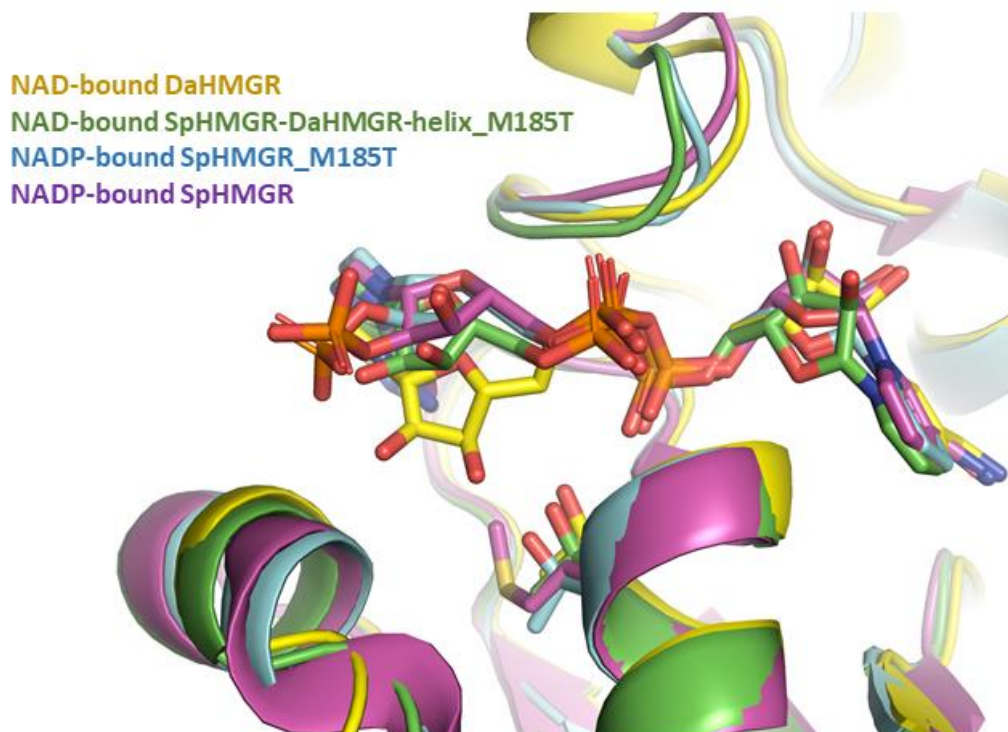
**Figure 15. Overlay of NAD-bound DaHMGR and NAD-bound DaHMGR\_T189M.** Overlay of NAD-bound DaHMGR (green) and NAD-bound DaHMGR\_T189M (cyan) shown in cartoon representation with T189 and M189 shown in sticks, respectively, and the cofactors in sticks colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring.



**Figure 16. Overlay of NAD-bound SpHMGR\_DaHMGR-helix and NADP-bound SpHMGR.** Overlay of NAD-bound SpHMGR\_DaHMGR-helix (cyan) and NADP-bound SpHMGR (grey) shown in cartoon representation the cofactor helices and cofactors in sticks colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring. D144 appears in alternate conformations as depicted.



**Figure 17. Overlay of NAD-bound SpHMGR\_DaHMGR-helix and NAD-bound DaHMGR.** Overlay of NAD-bound SpHMGR\_DaHMGR-helix (cyan) and NAD-bound DaHMGR (grey) shown in cartoon representation the cofactor helices and cofactors in sticks colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring. D144 appears in alternate conformations as depicted and a water molecule is represented by a red sphere. Interactions are depicted as dashed lines and colored based on interactions formed from its respective protein structure.



**Figure 18. Overlay of various DaHMGR and SpHMGR wildtype and mutant proteins showing the higher and lower binding of the cofactor.** Cartoon representation of wildtype DaHMGR, wildtype SpHMGR, and mutant SpHMGR structures colored according to the key. The cofactor and the corresponding methionine/threonine residue is shown in sticks and colored as follows: N colored blue, O colored red, P colored orange, C colored according to its cartoon coloring.

## 4.8. References

- [1] A. J. Covarrubias, R. Perrone, A. Grozio, and E. Verdin, “NAD<sup>+</sup> metabolism and its roles in cellular processes during ageing,” *Nat Rev Mol Cell Biol*, vol. 22, no. 2, pp. 119–141, Feb. 2021, doi: 10.1038/s41580-020-00313-x.
- [2] J. Zhang, N. Sonnenschein, T. P. B. Pihl, K. R. Pedersen, M. K. Jensen, and J. D. Keasling, “Engineering an NADPH/NADP<sup>+</sup> Redox Biosensor in Yeast,” *ACS Synth. Biol.*, vol. 5, no. 12, pp. 1546–1556, Dec. 2016, doi: 10.1021/acssynbio.6b00135.
- [3] W. Xiao, R.-S. Wang, D. E. Handy, and J. Loscalzo, “NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism,” *Antioxidants & Redox Signaling*, vol. 28, no. 3, pp. 251–272, Jan. 2018, doi: 10.1089/ars.2017.7216.
- [4] S. T. Rao and M. G. Rossmann, “Comparison of super-secondary structures in proteins,” *Journal of Molecular Biology*, vol. 76, no. 2, pp. 241–256, May 1973, doi: 10.1016/0022-2836(73)90388-4.
- [5] J. K. B. Cahn, C. A. Werlang, A. Baumschlager, S. Brinkmann-Chen, S. L. Mayo, and F. H. Arnold, “A General Tool for Engineering the NAD/NADP Cofactor Preference of Oxidoreductases,” *ACS Synth. Biol.*, vol. 6, no. 2, pp. 326–333, Feb. 2017, doi: 10.1021/acssynbio.6b00188.
- [6] C. R. Bellamacina, “The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins,” *FASEB j.*, vol. 10, no. 11, pp. 1257–1269, Sep. 1996, doi: 10.1096/fasebj.10.11.8836039.
- [7] J. A. Friesen, C. M. Lawrence, C. V. Stauffacher, and V. W. Rodwell, “Structural Determinants of Nucleotide Coenzyme Specificity in the Distinctive Dinucleotide Binding Fold of HMG-CoA Reductase from *Pseudomonas mevalonii* †,” *Biochemistry*, vol. 35, no. 37, pp. 11945–11950, Jan. 1996, doi: 10.1021/bi9609937.
- [8] E. R. Ragwan, E. Arai, and Y. Kung, “New Crystallographic Snapshots of Large Domain Movements in Bacterial 3-Hydroxy-3-methylglutaryl Coenzyme A Reductase,” *Biochemistry*, vol. 57, no. 39, pp. 5715–5725, Oct. 2018, doi: 10.1021/acs.biochem.8b00869.
- [9] B. H. Schwarz *et al.*, “Kinetic characterization of an oxidative, cooperative HMG-CoA reductase from *Burkholderia cenocepacia*,” *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1844, no. 2, pp. 457–464, Feb. 2014, doi: 10.1016/j.bbapap.2013.11.015.
- [10] M. Hedl *et al.*, “*Enterococcus faecalis* Acetoacetyl-Coenzyme A Thiolase/3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase, a Dual-Function Protein of Isopentenyl Diphosphate Biosynthesis,” *J Bacteriol*, vol. 184, no. 8, pp. 2116–2122, Apr. 2002, doi: 10.1128/JB.184.8.2116-2122.2002.

- [11] B. R. Miller and Y. Kung, “Structural Features and Domain Movements Controlling Substrate Binding and Cofactor Specificity in Class II HMG-CoA Reductase,” *Biochemistry*, vol. 57, no. 5, pp. 654–662, Feb. 2018, doi: 10.1021/acs.biochem.7b00999.
- [12] D.-Y. Kim, C. V. Stauffacher, and V. W. Rodwell, “Dual coenzyme specificity of *Archaeoglobus fulgidus* HMG-CoA reductase,” *Protein Sci.*, vol. 9, no. 6, pp. 1226–1234, 2000, doi: 10.1110/ps.9.6.1226.
- [13] A. E. Theivagt, E. N. Amanti, N. J. Beresford, L. Taberner, and J. A. Friesen, “Characterization of an HMG-CoA Reductase from *Listeria monocytogenes* That Exhibits Dual Coenzyme Specificity,” *Biochemistry*, vol. 45, no. 48, pp. 14397–14406, Dec. 2006, doi: 10.1021/bi0614636.
- [14] A. Beier *et al.*, “Switch in Cofactor Specificity of a Baeyer-Villiger Monooxygenase,” *ChemBioChem*, vol. 17, no. 24, pp. 2312–2315, Dec. 2016, doi: 10.1002/cbic.201600484.
- [15] M. Katzberg, N. Skorupa-Parachin, M.-F. Gorwa-Grauslund, and M. Bertau, “Engineering Cofactor Preference of Ketone Reducing Biocatalysts: A Mutagenesis Study on a  $\gamma$ -Diketone Reductase from the Yeast *Saccharomyces cerevisiae* Serving as an Example,” *IJMS*, vol. 11, no. 4, pp. 1735–1758, Apr. 2010, doi: 10.3390/ijms11041735.
- [16] A. Lerchner, A. Jarasch, and A. Skerra, “Engineering of alanine dehydrogenase from *Bacillus subtilis* for novel cofactor specificity: Engineering of *B. subtilis* AlaDH,” *Biotechnology and Applied Biochemistry*, vol. 63, no. 5, pp. 616–624, Sep. 2016, doi: 10.1002/bab.1414.
- [17] P. Bubner, M. Klimacek, and B. Nidetzky, “Structure-guided engineering of the coenzyme specificity of *Pseudomonas fluorescens* mannitol 2-dehydrogenase to enable efficient utilization of NAD(H) and NADP(H),” *FEBS Letters*, vol. 582, no. 2, pp. 233–237, Jan. 2008, doi: 10.1016/j.febslet.2007.12.008.
- [18] P. D. Adams *et al.*, “PHENIX : a comprehensive Python-based system for macromolecular structure solution,” *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. 2, pp. 213–221, Feb. 2010, doi: 10.1107/S0907444909052925.
- [19] A. M. Chánique and L. P. Parra, “Protein Engineering for Nicotinamide Coenzyme Specificity in Oxidoreductases: Attempts and Challenges,” *Front. Microbiol.*, vol. 9, p. 194, Feb. 2018, doi: 10.3389/fmicb.2018.00194.



## **Chapter V: Applications for HMGR in natural product, drug design, and metabolic engineering research**

### **5.1. Summary**

In Chapters Two through Four, experiments in several class II HMGRs were aimed at characterizing various biophysical characteristics, including solution-state studies (Chapter Two), characterizing the affinity of substrate and cofactor and related structural dynamics (Chapter Three), and modifying its cofactor preference (Chapter Three). Throughout these endeavors, there were discussions of how these characteristics manifest and their immediate significance in light of prior work on class II HMGRs. In this final Chapter, we utilize this knowledgebase to inform future avenues of interest by looking at specific experimental results and how they might contribute towards new research.

Three major avenues of research will be discussed: 1) HMGR's role in natural product synthesis, 2) metabolic engineering involving HMGR, and 3) drug development targeting HMGR. The specific experimental results that pertain to the structural basis of cofactor preference, oligomerization, and substrate affinity, will inform efforts at understanding and exploiting HMGR in natural product biosynthesis and drug discovery with specific focus given towards efforts at being able to fine-tune DaHMGR activity with respect to overall activity and cofactor utility. A paradigm that relates cofactor preference, reaction direction, oligomerization and the nature of products formed will be proposed. Unique ways of modulating or inhibiting HMGR activity which include unique hotspots at the active site as well as allosteric sites will be discussed. In sum, this closing chapter extrapolates the efforts of the previous sections and propels exciting areas of research that pertain to this biologically important enzyme.

## 5.2. Introduction

As pointed out in Chapter One, natural products continue to be an exciting reservoir of compounds that have benefit in a wide range of industries. It is, therefore, remarkable to consider that only a small fraction of the world's biodiversity has been evaluated for its potential, leaving a majority of the natural products available in nature open for discovery [1]. Equally exciting as the prospect of discovering new compounds from nature for society's benefit, is the research and development into the proteins that perform the necessary steps within the pathways that generate these compounds. By expanding our understanding of the way these proteins work there will be greater insight into the biosynthesis of these commodity compounds. Moreover, elucidating the mechanisms of how these proteins from different kingdoms end up producing often the exact same compound is an intriguing undertaking for the evolutionary biologist. Likewise, there is ripe opportunity for the structural biologist to discern the structural traits that result from these evolutionary processes which equips the synthetic biologist with the tools needed to engineer desired commodity compounds. The limiting step, therefore, is the research and development into these endeavors.

We begin this concluding chapter by first fixing HMGR as an important enzyme in the field of natural products. This will allow us to return to Chapter One and re-introduce an important observation: natural products are a promising field for producing compounds of interest. The implication of the current study expands our understanding of not only HMGR itself, but on HMGR-dependent reactions in natural product discovery. Secondly, following this discussion on HMGR and natural products, we consider metabolic engineering, and how the research done here can contribute towards HMGR-dependent production of natural products. Through insight gained into how HMGR cofactor specificity is governed and by being able to modulate its behavior, we

will explore how one can customize metabolic pathways in metabolic engineering efforts. Lastly, in the third section, we move away from natural products and metabolic engineering and look at HMGR itself as a potential drug target for antibiotics. Considering HMGRs from specific homologs on a molecular level, one can design inhibitors that are specific to a particular HMGR, and by extension, specific to the respective pathogen.

In sum, in this final chapter, with the knowledge gained from the research presented thus far, we begin to extrapolate how these efforts could be applied in natural products, metabolic engineering, and drug development and will be addressed as follows:

HMGR and natural products:

- ❖ *Why are natural products important?*
- ❖ *What is the significance of HMGR-dependent reactions in natural product research?*

HMGR and metabolic engineering:

- ❖ *What does metabolic engineering offer?*
- ❖ *How can we utilize intrinsic HMGR activity or modify it for synthetic biology efforts?*

HMGR and drug design:

- ❖ *What makes a good drug?*
- ❖ *How can a drug target class II HMGRs?*

### **5.3. HMGR and natural products**

#### **5.3.1. Why are natural products important?**

Throughout history, creating compounds for medicinal use has always been of great interest. Research and development into drug design have since spanned the globe motivated by the need to generate safe and life-sustaining medicine often with enormous financial incentives. When assessing the strategies of drug development, while in recent years a substantial amount of drugs are made from purely synthetic approaches, an even larger amount of drugs are inspired by what is already available in nature in the form of natural products [2]. As it pertains to drug development, these secondary metabolites can either have activity in its original form, are synthetically modified to an improved version, or they provide scaffolds for natural product models that offer medicinal benefits [1]. A popular example of a natural product derivative that was discovered to comprise medicinal use is the ubiquitous pain and anti-inflammation medicine, aspirin. Its active ingredient, acetylsalicylic acid, was synthesized based on the natural product, salicin, isolated from the bark of the willow tree *Salix alba* [1].

As a source for medicine, natural products offer the highest success rate in producing potential drug leads [1]. As pointed out by Newman and Cragg, in the almost four decades spanning Four Decades from 01/1981 to 09/2019, natural products have contributed, in some fashion, towards about half of all small molecules approved for medicinal purposes in the USA [3]. Figure 1 illustrates the proportionally sizable percentage of compounds that constitute drug sources implicated by natural products, demonstrating just how much influence these compounds have had on drug development. In their comprehensive review, Newman and Cragg give much attention to anticancer agents and infectious disease related drugs, with natural products featuring prominently among the approved medicine aiming to cure or treat these illnesses.

In addition to their roles in medicine, the use of natural products extends to other fields such as the production of biofuels, essential oils, nutrition and fragrances [4]. Such endeavors not

only prove beneficial to society but offer a lucrative industry to explore. For example, one estimate suggests that globally, the market for flavor and fragrances approaches ~40 billion dollars and cosmetics will reach over 800 billion dollars by 2023[5]. The demand for natural and sustainable products underscores the importance of finding efficient ways to produce these natural products [5]. Microbial production of natural products in these efforts has become increasingly popular [4]. Needless to say, the field of natural products, whether it be for the benefit of medicine or for the production of other chemicals of interest, is an attractive area of study and underscores the importance of research aiming to shed light on the biosynthesis of these important biomolecules.

### 5.3.2. What is the role of HMGR-dependent reactions in natural product research?

As pointed out in Chapter One, isoprenoids constitute the largest and most structurally diverse class of natural products. Despite its expansive presence across biology, the biosynthesis of isoprenoids is channeled through two major pathways, the DXP and mevalonate pathway. HMGR, is the first-committed step of one of these obligate pathways, thus fixing HMGR as a key enzyme across a substantial proportion natural product biosynthesis. As such, any investigation into HMGR contributes towards our growing understanding of how these important, often life-sustaining molecules, are made.

As it pertains to natural products, one important insight gained from this research involves looking at the preferred direction of HMGR activity. The forward direction is the reductive reaction, converting HMG-CoA to mevalonate. The reverse direction is the oxidative reaction. Knowing the preferred direction of an HMGR homolog has implications on the types of molecules that are made in its native context. As described in Chapter One, in the mevalonate pathway, when the forward reaction is considered, HMG-CoA is converted to mevalonate which goes on to

produce isopentyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), the building blocks for isoprenoid biosynthesis. However, this assumes that the forward reaction is preferred in its physiological context. An example of this is seen in archaea where the forward reaction leads to the production of isoprenoids that are important for the synthesis of, among other things, its membranes which are made up of glycerol ether linked to isoprenoid alcohols[6]. In contrast, HMGR from *Pseudomonas mevalonii* prefers to work in the reverse direction. Instead of using HMGR to produce the precursors for isoprenoids, it uses mevalonate as the sole source of carbon and energy, converting it to HMG-CoA to be eventually converted to acetyl-CoA, a central metabolite [6], [7]. Moreover, research into inhibitors that target HMGR activity would benefit from such knowledge since inhibiting HMGR activity also prevents the formation of HMGR-dependent downstream products. Therefore, the preferred direction of HMGR in its native context has major implications on the molecules being generated.

One way to assess preferred reaction directionality is to observe which cofactor is preferred. NADP(H) and NAD(H) often have different roles in the cell where NAD<sup>+</sup> is primarily used as an oxidant and NADPH as a reductant as discussed in Chapter Four [8]. This is revealed in low values, 0.001 to 0.01, of NADH/NAD<sup>+</sup> ratios in contrast to higher values, 15-60, in NADPH/NADP<sup>+</sup> ratios in various cellular environments[9]. In the context of HMGR, if an enzyme prefers NADP(H), it is likely that the forward reaction is preferred. For example, *Streptococcus pneumoniae* HMGR is an NADP(H)-dependent HMGR, and one study showed that the production of isopentyl pyrophosphate via the mevalonate pathway was essential for survival in *Streptococcus pneumoniae*, substantiating the claim that the forward reaction is preferred [10], [11]. In contrast, as mentioned above, PmHMGR, which prefers the reverse direction, is an NAD(H)-preferring HMGR [12]. Hence, with mounting evidence of this pattern, one could reasonably posit the types

of molecules HMGR predominantly generates, be it natural product or metabolic precursors, in its physiological context, by ascertaining its cofactor preference.

The research revealed in Chapter Three reveal DaHMGR as an NAD(H)-preferring HMGR. Using steady-state kinetics, we observe a significant preference for NAD(H). This preferred cofactor was subsequently used for crystallization experiments where NAD(H)-bound HMGR showed the important interactions that favor NAD(H). As highlighted, an important cofactor helix, which contains an aspartate (Asp146) coordinates the hydroxyl groups on the adenine-ribose moiety thus facilitating the cofactor preference towards NAD(H). This suggests that DaHMGR prefers the reverse direction and, therefore, will have implications on the physiological relevance of mevalonate acquisition and utility in the cell as well as HMG-CoA production.

A related phenomenon to which this research contributes towards involves the relationship between oligomerization and cofactor preference. There appears to be a developing paradigm within class II HMGRs where HMGRs that prefer NADP(H) as its cofactor manifest as a dimer whereas HMGRs that prefer NAD(H) are able to form hexamers. The work in Chapter Two agrees with this paradigm showing that NADPH-preferring SpHMGR exists as a dimer in solution as revealed by Size Exclusion Chromatography, with no higher order oligomerization evident. A predominately single peak in the expected dimer size is observed. This is also reflected in the crystal structure of SpHMGR where SpHMGR is revealed as a homodimer in the crystal lattice. In contrast, the crystal lattice of the crystal structure of DaHMGR can create a hexamer in the crystal lattice. Upon further solution-state analysis performed using Size Exclusion Chromatography and Small-Angle X-ray Scattering, DaHMGR exists in equilibrium between a hexamer and dimer. Its ability to form the hexameric assembly is noteworthy in the context of

oligomerization and its respective cofactor preference. Thus, the work here further supports the developing paradigm that NADP(H)-preferring HMGRs and NAD(H)-preferring organisms exist in different oligomeric states in its native environments and more specifically points to the ability of NAD(H)-preferring HMGRs to form higher order oligomerization.

In the context of natural product biosynthesis, the ramifications of these observations are noteworthy. Further work needs to be done in this regard, but so far, it appears that cofactor preference, preferred reaction direction in its respective cell, and oligomerization are related. Their does not appear to be causative influences between cofactor preference and oligomerization, but the correlation exists. Therefore, it may be possible to look at any given one of these aspects, in an uncharacterized HMGR, and predict the remaining unknowns. For example, it may become possible to simply look at the sequence of the cofactor helix, predict its cofactor preference based on the structural residues that constitute the structural determinants that govern cofactor specificity, and thereby classify its preferred directionality, its oligomerization preference, and ultimately the types of molecules it generates in the cell. One could assess whether HMG-CoA or whether mevalonate are preferred products which has major implications on the types of downstream products that can be generated in the context of natural product biosynthesis.

For example, in the case of *Staphylococcus aureus*, based on the observation that mevalonate production is required for survival[13], the HMGR reaction probably favors the forward direction, converting HMG-CoA to mevalonate. The cofactor helix, made up of residues AYPSIKAR, lacks the conserved aspartate that would coordinate the 2'-hydroxyl of NAD(H) and instead has the bulkier tyrosine that helps to sterically position the phosphate of NADP(H) to coordinate the charged residues found later in the cofactor helix. This suggests that it would be NADP(H)-dependent which agrees with the above assumption that the preferred direction is the



reductive reaction. However, the oligomeric state of HMGR remains unknown. Using the developing model proposed here, we can predict it to be a dimer. As such, the work done here has the potential to inform aspects of individual HMGRs as well as class II HMGRs in general and offers insight into HMGR-dependent natural product biosynthesis.

## **5.4. HMGR and metabolic engineering**

### 5.4.1. What does metabolic engineering offer?

Metabolic engineering provides the opportunity for researchers to investigate methods of producing a wide range of compounds with uses as fuels, medicine, and other commodity chemicals spanning both commercial and health-related industries. These methods involve harnessing or modifying the metabolism of a host in order to leverage the production of native or non-native metabolites[14]. These techniques not only offer ways to produce novel compounds for use in new ways, but also offer alternative ways of producing existing chemicals. Earlier in this section, we highlighted the importance of natural products in producing chemicals of interest. Thus, it comes as no surprise that metabolic engineering of natural products is revealing itself as a promising approach to acquiring compounds for benefit to society [15].

To this end, there are efforts towards producing compounds such as fossil fuels that are in jeopardy of being depleted, or greener alternatives in response to climate change and environmental pollution [16]. These methods can involve using cell-based systems where, for example, living microbes are modified to produce desired compounds. They can also involve cell-

free systems where specific proteins are isolated and incorporated into a synthetic pathway that leads to the production of desired compounds.

In a cell-based biosynthetic platform, a common host used in metabolic engineering is the bacterium *Escherichia coli* (*E. coli*). Since *E. coli* is such a malleable and efficient host, biosynthetic pathways found elsewhere in nature can be inserted into *E. coli*, allowing the abundant production of desired products. As depicted in Figure 2, the exogenous pathway becomes incorporated into the machinery of the host cell in a way that does not drastically disrupt the survival of the host and is able to perform targeted reactions to produce highly specific products. In one study, researchers were able to generate a modified strain of *E.coli* that could produce a class of branched-chain fatty acids that constitute precursors for renewal fuels[17]. Normally, *E.coli* does not naturally produce branched chain fatty acids, however by customizing the supply of branched-chain amino acids in the growth medium and by modifying the amino acid biosynthesis pathways, which included upregulating certain enzymes, they were able to generate branched-chain fatty acids [17].

There are, however, limitations to using cell-based systems which revolve around the need to maintain homeostasis within the host that prevents adequate product formation. For example, microbes are usually only able to perform optimally at certain temperatures, or the implicated pathways end up creating toxic amounts of products or intermediates [16]. Thus, one alternative relies on a cell-free system where enzymes are isolated, combined, and incorporated into a synthetic pathway, *in vitro*, to generate a desired product. In one study, researchers were able to design an entirely artificial glycolytic pathway that converts glucose to pyruvate using only four enzymes [16]. The pathway, made of enzymes from a variety of different source organisms, eliminated phosphorylation and dephosphorylation steps required in typical glycolytic pathways,

and used only NAD<sup>+</sup> as a molecular shuttle. The researchers took this concept further and utilized two enzymes to convert pyruvate to ethanol, and five enzymes to convert pyruvate to isobutanol. Of importance, is that in order to be selective towards a specific intermediate in the pathway, one of the enzymes, NADPH-dependent aldehyde dehydrogenase, was engineered to have greater affinity for NADH since the pathway already utilized NAD(H) [16].

In both the above approaches, cell-based and cell-free biosynthetic platforms, a more complete understanding of the pathways that are implicated in product formation, as well intimate knowledge of the individual enzymes involved, are prerequisites of being able to utilize these platforms effectively. For example, in cell-based systems, good understanding of the aspects that limit and regulate the pathway, the types of intermediates that are generated, and how enzymes exist in its cellular environment, are vital. In a cell-free system, it will be advantageous to know the kinetic parameters such as catalytic efficiency, specific activity, and substrate affinity of enzymes within the synthetic pathway to be able to predict and control the flow of product formation. It would also be highly beneficial to have available the means to modify the behavior of proteins in a predictable manner as was the case of the cell-free system discussed above.

#### 5.4.2. How can we exploit intrinsic HMGR activity?

##### 5.4.2.1. The feasibility of exploiting HMGR activity

There have already been numerous cell-based and cell-free platforms that have exploited HMGR activity. The research from this study adds to and extends our understanding of some of the ways in which HMGR has been and can be used. For example, in the cell-based biosynthetic study where researchers developed an *E.coli* strain to express the mevalonate pathway from *Saccharomyces cerevisiae*, they interchanged the gene encoding HMGR with those from various

organisms and found that when HMGR from *Delftia acidovorans* (DaHMGR) was used, there was higher product formation, amorphadiene, the antimalaria precursor [18]. In Chapter Three, DaHMGR was kinetically characterized and, as pointed out in Table 1, DaHMGR has the lowest  $K_M$  out of various class II HMGRs. This suggests that DaHMGR has a comparably higher affinity for HMG-CoA than other HMGRs and might account for the reason DaHMGR was the most effective at producing the most amorphadiene.

However, this interpretation is complicated by the observation that in the study mentioned above, *Staphylococcus aureus* HMGR (SaHMGR) showed the highest mevalonate-producing rate –the forward reaction –and DaHMGR had the highest HMG-CoA producing rate –the reverse reaction –when assessed *in vitro* with their preferred cofactor and *in vivo* under aerobic conditions. Under anaerobic conditions, this pattern was reversed revealing higher mevalonate production using DaHMGR than SaHMGR. As the authors point out, concentrations of NADH in *E. coli* are higher under anaerobic conditions. Thus, the work in Chapter Three which characterizes DaHMGR as an NADH-specific HMGR aids in our interpretation of this phenomenon. As depicted in Chapter Three, DaHMGR has significantly lower activity when using NADPH and strongly prefers NADH. Of note is that DaHMGR also has the lowest  $K_M$  for NADH of the homologs represented in Table 1. Therefore, under anaerobic conditions, with higher concentrations of NADH present in *E. coli*, DaHMGR which has a high affinity to HMG-CoA and prefers NADH, can perform the forward reaction producing HMG-CoA at high rates.

This does not, however, explain why DaHMGR eventually went on to produce more amorphadiene since the experiment relied on the forward reaction under aerobic conditions. HMGR is often considered the rate-limiting step, and as such, under aerobic conditions, one would expect SaHMGR, which had the highest forward reaction rate under aerobic conditions, to be the

most efficient enzyme to produce larger quantities of amorphaadiene. Further investigation into this apparent contradiction led the researchers to observe the amount of unused mevalonate present during *in vivo* production of amorphaadiene and determined that there was unused mevalonate in extracts from SaHMGR experiments. This suggested to them that, while SaHMGR was effective at making high amounts of mevalonate, the subsequent enzyme in the mevalonate pathway, namely mevalonate kinase, which converts mevalonate to mevalonate-5-phosphate, was being inhibited by high concentrations of mevalonate.

Now that DaHMGR has been studied more rigorously in the present research, the work done here provides a quantitative characterization of how it may function *in vivo*. More specifically, one could look at the kinetic parameters and enable a quantitative approach at metabolic engineering geared towards HMGR-dependent reactions. Beyond explaining and understanding, metabolic engineering aims to modify behavior for desired outcomes.

#### 5.4.2.2. Modifying HMGR activity

In the study above, when trying to explain how DaHMGR is able to produce more amorphaadiene than when other HMGRs were used, the authors proposed that DaHMGR is able to balance HMG-CoA levels below the threshold that would lead to cellular toxicity while also balancing mevalonate levels below concentrations that would inhibit mevalonate kinase activity [18]. This balancing act is directly related to the affinity of substrates and products as well as activity, where fine-tuning allows for optimal function in the specific pathway it is incorporated in to. This fine-tuning of DaHMGR activity can be further exploited and can be investigated *in vitro* by modulating its activity. This modulation can involve enhancing affinity, with the hope of enhancing catalytic efficiency, or it can involve attenuating affinity resulting in diminished activity. Especially in the case of attenuating binding, one must be careful not to totally destroy

affinity and instead make finer adjustments that lead to controlled modulation of activity. Thus, a highly nuanced approach, where intimate knowledge of the enzyme dynamics, is warranted. The hope is that one can fine tune, in a controlled manner, the various factors, such as substrate and product affinity and enzyme activity, for optimal product formation.

One approach to modulating activity would be to examine the importance of higher order oligomerization. As mentioned in Chapter Two, DaHMGR exists as a hexamer made of trimers of homodimers. In light of the fact that only the homodimer is required for activity in many class I and class II HMGRs, it is likely that homodimers of DaHMGR would still be active. And since, as pointed out in Chapter Two, oligomerization often has implications on the activity of the protein which is most commonly observed by revealing higher activity with higher order oligomerization, one could potentially modulate DaHMGR activity by enhancing or preventing the formation of the hexamer. This can be achieved by exploring the hexameric interface and homing in on residues that can be mutated. These mutations could be aimed at furthering the interactions that stabilize this interface or they can disrupt of the formation of the hexamer. Equipped with the structures of DaHMGR and the biophysical approaches investigated in Chapter Two, the prospect of this approach is promising.

To disrupt the hexameric interface of DaHMGR, there are several residues already being investigated in the Kung Lab. Glu308 and Gln133 are examples of residues that can be targeted (Ashley Yang and Yan Kung, unpublished). Glu308 makes interactions with the backbone His125 and Gly126 of a neighboring molecule. Gln133 interacts with the backbone of Thr105, and with the sidechain of Q104. As pointed out in Chapter Two, Peacock et al. suggest that DaHMGR also has a potential salt bridge between Glu320 and the side chain of Arg128, and another potential salt bridge exists between Asp123 and Arg226 [19]. In all the instances, mutating glutamate,

glutamine, aspartate or arginine, to alanine will prevent these interactions. In addition, at the hexameric interface, the helix containing residues Gln133 to His125 spans across a series of anti-parallel B-sheets. By mutating Gly126 to a glutamine, it is possible that the large side chain introduced through this mutation would prevent the helix from interfacing with the B-sheets (Ashley Yang and Yan Kung, unpublished.) Figure 3 shows images of these various interactions at this interface. Towards the goal of disrupting the hexameric interface, it seems highly likely that a combination of a variety of the above mutations would prove successful to this end. Size Exclusion Chromatography can be used to characterize the oligomeric state of the expressed proteins, and kinetic analysis can be performed to assess the resulting catalytic activity.

In addition, Peacock et. al., analyzed the interfaces of the three NAD-preferring class II HMGRs whose structures have been deposited in the Protein Data Bank, all of which appear to form hexamers, PmHMGR, DaHMGR and BcHMGR, and they note that the hexameric interface of DaHMGR is the strongest with two potential salt bridges and eight hydrogen bonds at the DaHMGR interface, one salt bridge and seven hydrogen bonds at the PmHMGR interface, and one salt bridge and two hydrogen bonds at the BcHMGR interface [19]. With such information, one could vary the strength of the interface in a rational way to either increase or decrease the strength of the interactions. In doing so, one could influence the equilibrium of oligomerization in solution such that increasing or decreasing the number and strength of these interactions could lead to greater or lesser proclivity to form hexamers. It remains to be determined, however, if and how oligomerization influences activity, but the crystal structures provided in this study, in addition to the structures already available in the Protein Data Bank, provide ample room for metabolic engineering efforts.

Another metabolic engineering approach of modulating activity would be to increase or decrease the affinity of substrate, intermediates, and products, thus influencing the rate at which the forward or reverse reaction occurs. An overlay of apo DaHMGR and HMG-CoA bound SpHMGR show a highly conserved binding pocket. However, one interesting difference exists at the solvent exposed region of HMG-CoA. A serine instead of an arginine flanks the adenine portion of HMG-CoA. In HMG-CoA bound PmHMGR, this arginine interacts with the adenine moiety through cation-pi stacking interactions as shown in Figure 4, similarly observed in CoA-bound-BcHMGR [19].

Further analysis into the sequence alignment of various class II HMGRs reveal that this residue is either aromatic –tyrosine or histidine—or arginine, but in SpHMGR this residue is a serine. Figure 4C shows a sequence alignment at this region. It is probable, that the aromatic side chains also undergo pi-pi stacking interactions between the ring structures and the adenine moiety in a similar manner as arginine where it is observed in the case of HMG-CoA-bound PmHMGR. Consequently, one could assess if and how these residues, namely tyrosine, histidine, arginine and serine, influence HMG-CoA affinity, and could rank these mutations in order of its influence on activity. Using this unique site is advantageous because it is likely that mutating other residues within the binding pocket would result in more drastic changes to HMG-CoA affinity as is emphasized by the highly conserved architecture. In contrast, this site already has some variability, and affords flexibility in metabolic engineering to be able to fine tune substrate affinity.

#### 5.4.2.3. Modulating HMGR cofactor preference

The entire of Chapter Four is dedicated to HMGR cofactor preference, and as pointed out, being able to control cofactor affinity and preference is valuable in the metabolic engineering arena. To this end, researchers could either enhance or diminish the affinity of a cofactor to an



HMGR, make an HMGR more promiscuous towards NAD(H) and NADP(H), or switch the cofactor preferences entirely. In the previous section, the importance of cofactor requirements within the context of the forward or reverse reaction within aerobic or anerobic conditions was proposed and emphasizes the importance of being able to control this feature.

The results of Chapter Four set forth a promising path towards modulating cofactor preference not only in DaHMGR and SpHMGR, which were the HMGRs investigated in that chapter, but also across class II HMGRs. The patterns observed and utilized, namely the residues that constitute the cofactor helix that is believed to dictate cofactor preference, is a potentially translatable approach where this paradigm could inform a variety of engineering efforts in HMGRs. However, as also discussed in Chapter Four, while there are some approaches that are translatable, it may be necessary to take each HMGR on a case-by-case basis.

In the case of NADH-preferring DaHMGR, for example, one could compare its cofactor helix, RDKVLIGL, with that of another NADH-preferring HMGR, PmHMGR, which has the cofactor helix of KDQLLSL in an attempt to investigate which helix would provide better cofactor affinity. From this scaffold, additional modifications to the cofactor helix could be made to influence cofactor affinity. It is conceivable that one could create a highly efficient DaHMGR based on its engineered ability to use very small amounts of NAD(H) which would be useful in a biosynthetic platform that needs to optimize cofactor requirements. If, on the other hand, one desires to diminish activity, to prevent the reaction from producing toxic amounts of the product, one could alter the cofactor helix accordingly.

In a biosynthetic platform that has both NAD(H) and NADP(H) it may also be advantageous to be able to modify DaHMGR with the ability to use both cofactors. This would be particularly important in a platform where the availability of either of these cofactors is of

importance. Chapter Four revealed a chimeric DaHMGR, DaHMGR\_SpHMGR-helix, which is able to use both NADH and NADPH comparably. While this chimeric construct has a slightly higher preference for NADP(H), it is significantly more promiscuous than wildtype DaHMGR. If, however, one desires to switch the cofactor preference of DaHMGR entirely, one could modify DaHMGR\_SpHMGR-helix with the addition of the T189M mutation discussed in Chapter Four. This chimeric DaHMGR, DaHMGR\_SpHMGR-helix+T189M, has a new preference for NADP(H) and can be used in a pathway where NADP(H) is the most abundant cofactor available. While work still needs to be done to improve the overall efficiency of the chimeric enzymes mentioned above, it is exciting to consider the successful proof-of-concept already demonstrated in the research presented here.

One could make the argument, that instead of engineering an existing HMGR to have different activity, whether it be for the purpose of exploiting oligomerization dynamics, substrate binding, or cofactor affinity, that one could simply use a different HMGR homology instead. However, this is not always a feasible option. As pointed out in the study that began this section, a variety of factors need to be considered when interchanging HMGRs. Examples of some of these factors include cofactor preference, its rates for the forward or reverse reactions, and how the HMGR performs under aerobic versus anerobic conditions. Ultimately balancing of all these factors in the context of the specific pathway would benefit from being able to customize them. For example, in the amorphadiene study, SaHMGR had the highest forward reaction rate, producing the most mevalonate under aerobic conditions with its preferred cofactor. This would suggest that SaHMGR would be the ideal HMGR to use. However, as already pointed out, DaHMGR was able to balance the aforementioned factors that resulted in the more effective amorphadiene production. Thus, while the intrinsic activity of each HMGR is valuable, the

importance of rationale-based engineering of HMGRs to be able to fine tune a particular HMGR for desired outcomes is a noteworthy pursuit which makes the work presented here valuable.

## **5.5. HMGR and drug design**

### **5.5.1. What makes a good drug?**

In a general, yet concise description, a good drug needs to be safe and effective; it needs to be safe with minimal side-effects, and effective at feasible dosages. In contrast to its concise aspirations, the complexity of achieving these standards is immense and requires a good understanding of parameters such as bioavailability, drug affinity, drug potency, and efficacy. Bioavailability refers to the amount of drug present in the blood supply after the drug is administered and is represented as a fraction of the original amount given [20]. Drug affinity is a measure of how strongly a drug binds its target and is often expressed in terms of the dissociation constant  $K_d$ , the concentration of one binding partner (A) required for 50% of the other binding partner (B) to become bound in an AB complex [20]. A related term is  $IC_{50}$  which is a measure of the concentration of a drug inhibitor that prevents 50% of the native substrate from being bound to the target [20]. Efficacy is a measure of the maximum possible effect of the drug in its physiological context, and potency, which is noted in terms of  $EC_{50}$ , is the concentration of drug required to produce 50% of the maximum effect [20].

A majority of drug compounds target proteins, in the form of receptors, enzymes and transport proteins, as a means of performing its therapeutic outcome [20]. In this context, good drug design requires interrogation of the drug-protein complex such that the quantity, position, and

strength of the interactions can all be modulated with flexibility which is a result of intimate knowledge of the structure and function of the protein. Typically, a high-throughput screen is used to identify lead compounds with a crude  $K_d$  to measure affinity of the drug to the target. From this scaffold, additional functional groups are removed, modified, or added, to increase the drugs therapeutic outcome.

What makes drug development increasingly complex, however, is that some of the above parameters –bioavailability, affinity, efficacy, and potency –can be mutually exclusive. For example, one could discover a compound that binds very tightly to the target protein, but it is possible that due to the makeup of the compound, the drug is unable to pass through the various metabolic routes to enter the blood at high enough doses. Thus, if its bioavailability is poor, then the overall effect of the compound is severely hampered. In this example, researchers could consider modifying the lead compound, by removing, modifying, or adding functional groups that aid in the drug's bioavailability. There is also the concern that the drug could interact with other proteins in the body causing toxicity issues. Again, the strategy could be to introduce structural features onto the compound that retain its binding affinity to the original target while disrupting its binding to unwanted alternative targets. As such, it is insufficient to simply screen for compounds that have high affinity to its target. Instead, it requires substantial downstream optimization to balance strong binding of the drug to the target as well as other factors that ultimately contribute towards its safety and effectiveness.

The types of drug-target interactions pertinent to drug design are as follows: electrostatic interactions, which are the strongest of interactions applicable to drug design, and are notably inversely proportional to distance and highly dependent on the nature of the environment; hydrogen bonds vary in strength and is dependent on the angle at which the electron-rich and

electron-deficient atoms interact; Van der Waals interactions are considered weak, highly distance-dependent interactions, but in greater quantity can be crucial in binding, such as those between a carbon skeleton of a drug and parts the enzyme; other interactions include dipole-dipole, ion-dipole and repulsive interactions[20]. Armed with high resolution crystal structures of the target protein, a researcher can utilize the above knowledge of the quantity, position, and strength of the interactions, in developing a drug that is finely tuned to exploit these interfaces while also balancing factors such as bioavailability and minimizing non-specific binding that collectively contributes towards favorable therapeutic outcomes.

In addition to understanding the drug-target complex, further knowledge about the protein's manifestation within the cell and body need to be considered. For example, there needs to be a good understanding of the relative expression and abundance of the protein within the cell and localization of the target protein in various parts of the host[21]. Furthermore, the way a protein presents itself in the cell, as compared to *in vitro*, is an important consideration. As outlined in Chapter Two, the way a protein presents itself in its native context can often be studied using solution-state techniques and include topics such as its dynamic three-dimensional constitution and its oligomerization states. Additionally, crystal structures of the target protein are vital in drug design efforts. Therefore, while substantial research and development is emphasized during the later drug approval processes, namely pre-clinical to approval phases, drug research relies on the work of basic research in studying the structure and function of the protein in isolation[22].

### 5.5.2. How can we target class II HMGRs at the active site?

#### 5.5.2.1. The feasibility of targeting class II HMGRs

It is not within the scope of this study to pursue drug design. On a drug research and development spectrum ranging from a concept to FDA approval, the data from the research presented here humbly positions itself at the earliest of stages. However, the results provided here, namely the solution-state studies in Chapter Two, and the relatively high-resolution structural data and kinetic parameters revealed in Chapter Three, provide a feasible foundation of basic research from which to work from. The overall aim would be to identify HMGRs from specific pathogens and observe if inhibiting HMGR from those pathogens would result in depriving the organism of necessary metabolites required for survival. This concept has already been explored as was demonstrated in one study where researchers showed that HMGR was essential for survival in *Staphylococcus aureus* by genetically perturbing the gene that encodes HMGR [13]. An inhibitor designed to prevent HMGR activity in *Staphylococcus aureus*, a bacteria that can sometimes cause serious infections with the concern of growing multi-drug resistant strains [23], could lead to the development of a novel antibiotic.

An important observation worth noting is that class I inhibitors, namely statins, are highly effective at disrupting class I HMGR activity which serves as the basis for modulating cholesterol synthesis as pointed out in Chapter One. However, statins do not bind to class II HMGRs as well as class I HMGRs with one study showing over  $10^4$ -fold higher  $K_i$  values for Lovastatin inhibiting class II HMGR over class I HMGR[24]. There are specific interactions between the drug and the residues of class I HMGRs that are lacking in class II HMGR which contribute towards the differences in inhibition profiles [24]. Of importance is that class I and class II HMGRs reveal changes in the residues that are crucial for catalytic activity as well as in regions that promote oligomerization as elaborated in Chapter Two[25].

Herein lies an important observation: the above data suggests that there are substantive differences between the active site architecture between both classes of HMGR. So, it is feasible, and in the case of statins, already shown that a drug can be specific to either class I or class II HMGR. This is advantageous because if the purpose of inhibiting HMGR is to disrupt HMGR from a pathogen, it would necessarily want to avoid targeting the host HMGR. Thus, the differences between the three-dimensional structural determinants that result in the varying HMGR inhibition profiles between classes of HMGRs aids in the feasibility of being able to select between the host HMGR and pathogenic HMGR.

To date, there lacks a comprehensive comparison of the functional differences between various homologs with respect to inhibition. Most studies assess inhibition with regards to a specific HMGR without comparing inhibition across HMGRs. There are, however, kinetic studies done using HMG-CoA for a variety of class II HMGRs. In Table 1, the  $K_M$  values of various class II HMGRs are shown in increasing order. These values range from 7.34 to 175  $\mu\text{M}$  HMG-CoA. This variation reveals that in some instances, the differences in  $K_M$  are small, for example, 19.8 versus 20  $\mu\text{M}$  HMG-CoA for LmHMGR and EfHMGR respectively. However, differences between DaHMGR and SpHMGR, for example, differ by 10-fold with values of 7.34 versus 75.86  $\mu\text{M}$  HMG-CoA.

This would suggest that there might be structural differences not only between classes of HMGR but also within class II HMGR that may facilitate drug design that is selective towards specific pathogens. One could interpret the kinetic values above as reason to suggest that ultimately drug design needs to be taken on a case-by-case basis considering that many of the  $K_M$  values vary only by 3-6-fold. Working with the exact structural and functional characteristics of a particular HMGR is warranted.

#### 5.5.2.2.Targeting HMGR at the active site

The crystal structures of DaHMGR described in Chapters Two and Three represent the first crystal structures of HMGR from *Delftia acidovorans*. In addition, at the time of publication, there were crystal structures of only two other class II HMGR homologs, namely PmHMGR and SpHMGR, and as such the publication of DaHMGR structures increased our ability to ascertain differences and similarities between HMGR homologs. Since then, the structure of BcHMGR has also been published. In addition to DaHMGR, high resolution structures of SpHMGR and SpHMGR mutant constructs were presented in this research. While structures of SpHMGR already exist, obtaining higher resolution structures, with novel insight, enhances our understanding on a molecular level. However, considering how ubiquitous HMGR is in biology, the need for additional crystal structures of various HMGR homologs still persists.

The catalytically important residues at the active site of class II HMGRs are highly conserved among class II HMGRs. Figure 5, an overlay of the active site of several class II HMGRs, illustrates how well conserved the residues and their respective positioning are. As such, it would be particularly difficult to design a drug that differentiates between species using rationale-based structural analysis. In practice, a highly effective method of finding compounds that bind uniquely to the active site is to use a high-throughput screen, and specifically, a high-throughput screen that reveals drugs that not only bind the active site but also inhibit the function of the enzyme. From this screen, compounds can be further investigated. To this end, there have already been some efforts which have revealed important findings that would facilitate future drug design that will be discussed here.

Prior investigations have suggested that a drug that targets class II HMGRs must include the -CH<sub>3</sub> and -OH groups on the  $\beta$ -carbon and the carboxyl group of mevalonate[26], [27], [24].



The importance of these functional groups makes sense considering that the substrates, intermediates, and products of this reaction, namely HMG-CoA, mevaldyl-CoA, mevaldehyde and mevalonate, all contain these moieties. It was shown that exchanging these functional groups with others resulted in a loss of inhibition[24]. The work presented in Chapters Three and Four, containing several structures of DaHMGR and SpHMGR with citrate bound in the active site, directly relates to this phenomenon. Citrate was introduced during crystallization experiments by using crystallization solutions with citrate as a buffering component. Citrate contains the -OH on the  $\beta$ -carbon and the carboxyl group of mevalonate, however it has a -CO<sub>2</sub> group instead of a -CH<sub>3</sub> on the  $\beta$ -carbon of mevalonate. This seemingly small difference in functional groups showed a correspondingly low citrate inhibition with a  $K_i$  of 170 mM. This inhibition is weak despite showing that citrate competitively binds with respect to the substrate by occupying the mevalonate binding site with some of the expected substrate interactions. This further emphasizes the importance of each of the above functional groups. If, for example, high-throughput screening of drug candidates reveal new compounds that lack some of the above structural characteristics, the initial drug lead can be modified to include these functional groups to enhance drug-HMGR interactions.

In Chapters Two and Three, there was an analysis of the various conformations the C-terminal domain (CTD) assumes at various points in the mechanism. Of importance is that the CTD interfaces with the rest of the protein differently through these various conformations such that different residues are presented to the active site at various points in the mechanism. For example, in the ternary structure of PmHMGR bound to HMG-CoA and NAD<sup>+</sup> where the CTD adopts a closed conformation, His381, Leu372 and Ile377 all of which are from the CTD, are implicated near the active site and proposed to be catalytically important [28]. His381 is the

catalytically important histidine described in Chapter Three. It also proposed that Leu372 and Ile377, form a part of a hydrophobic pocket along the inner face of the first helix of the CTD which may help to shield the hydride transfer reaction from solvent and create a favorable hydrophobic environment for the nicotinamide moiety[28]. In the HMG-CoA-bound SpHMGR structure which positions the CTD in a partially closed conformation, Lys380 and Lys384, both of which are from the CTD, directly interact with the diphosphate moiety of HMG-CoA. In both above instances, the position of the CTD introduces new interactions at the active site along the substrate-binding motif that provides additional hot spots to target which are beyond the conserved active site residues. For example, the hydrophobic residues offer the medicinal chemist the opportunity to utilize Van der Waals interactions using carbon skeletons as previously mentioned.

In this research, novel locations of the CTD are revealed. One structure of DaHMGR reveals the CTD in a flipped conformation with citrate bound in the active site. Citrate mimics the binding of mevalonate and the HMG moiety of HMG-CoA. While there are no specific interactions between residues of this CTD in the flipped conformation and citrate, the second helix, instead of the first helix mentioned above, lays adjacent to the active site. In this orientation, residues of the second helix can be exploited as drug targets. One example is Arg405 which is in reasonable proximity to the substrate and cofactor. Figure 6 shows the positioning of the second helix and the location of the Arg405 with respect to the substrate and cofactor. Arginine offers the medicinal chemist the opportunity to create ionic interactions which are the strongest drug-protein interactions as previously mentioned.

Targeting specific HMGRs will be effective using high-throughput screening of small molecules. After this initial phase of finding promising leads, rationale-based optimization of drug design will be enhanced by our understanding of the structural insight obtained from the

aforementioned structures. New insight obtained from relatively high-resolution structures of DaHMGR and SpHMGR with regards to interactions at the active site and substrate binding sites offer additional opportunity to strengthen drug-protein interactions. These additional hot spots are important when considering how similar the active sites between HMGRs are and may be necessary when trying to design drugs that are highly specific.

### 5.5.2.3. Targeting allosteric sites of HMGR

In the previous section, the CTD was discussed in the context of generating drugs that target the residues that exist on the CTD that interface with the active site. The goal was to design substrate inhibitors that exploit CTD interactions near the active site. However, in this section, we continue exploring the role of the CTD by proposing ways to target the CTD at other locations. As pointed out in Chapter Three, the CTD is required for catalysis, so it seems plausible that designing a drug that prevents the activity of the CTD offers an opportunity to halt the entire reaction by allosterically regulating the enzyme.

The CTD plays at least two important roles. First the role is to position the catalytically important histidine at the active site at a precise step in the reaction mechanism. The second role is to allow substrates, intermediates, and products to enter, be retained, or leave, at the appropriate step within the reaction mechanism. By performing its catalytically important duties, the CTD traverses through various conformations, including being in an open conformation that may exist like an extended arm that is able to move freely, to then interfacing with the rest of the protein in partially closed, closed, and flipped conformations. It is possible, then, to create a small molecule that prevents the CTD from being able to adopt one or more of the catalytically important conformations. As expected, HMGR mutants where the CTD was completely truncated showed no activity (Edwin Ragwan and Yan Kung, unpublished).

There are several key positions on the CTD that would be good targets and involve the two important roles previously mentioned. One target would be the catalytically important histidine. The structures of the various conformations of the CTD presented in this research emphasize the inherent flexibility of the CTD. In the DaHMGR structure containing the CTD in the flipped conformation, the catalytically important histidine points away from the protein and are exposed (see Figure 7). Furthermore, the open conformations of the CTD in SpHMGR discussed in Chapter Four shows that there are multiple possible open conformations. Thus, in an open conformation, the CTD is vulnerable to being targeted by small molecules at various locations along the peptide chain. Therefore, either in the flipped conformation or in the open conformations, a small molecule, such as a peptide fragment, designed to target the catalytically important histidine and flanking residues could function as an effective antagonist to overall protein activity.

Additional targets of the CTD include the regions that interface with the protein. By generating compounds that prevents the closure of the CTD one could kill both roles of the CTD, namely the positioning of the histidine at the active site and the controlling the flux of substrates, intermediates, and products. Prior to this research, these interfaces were limited to closed and partially closed conformations, but the novel flipped conformation revealed through structural analysis of DaHMGR provides additional approaches. A related target would be the flexible linker that joins the CTD with the rest of the protein. A peptide fragment designed to immobilize this linker, or one that would sufficiently disrupt its flexibility, will prevent the overall movement of the CTD.

Yet another allosteric target would involve preventing oligomerization of HMGR. HMGR is catalytically active as a dimer where the substrate of one monomer interacts with the cofactor of its adjacent monomer. The dimerization interface, as discussed in Chapter Two, is extensive and

well conserved, and requires a well-orchestrated interlocking between the N-terminal domains of each monomer. It is unknown how and when this dimerization event takes place. It does, however, open the possibility of designing a small molecule that would prevent this dimer from forming. Moreover, solution-state studies of DaHMGR revealed in Chapter Two suggests that it exists as a hexamer made up of a trimer of dimers. It remains to be investigated how the activity of the hexameric structure compares to a dimer. If the hexamer is more catalytically active, one could target this higher order oligomerization motif.

Most drugs that work on proteins directly bind to the active site however this allows for the drug to bind to the active sites of proteins that have related functionality which results in non-specific interactions leading to side effects [29]. A marked advantage of targeting allosteric sites, therefore, is the ability to be highly selective towards regions of a protein that do not necessarily share conserved motifs, thus allosteric regulation offers a new paradigm for drug discovery [30]. In the case of HMGR, while the catalytic histidine is well conserved, the residues flanking the histidine are not. Therefore, generating a peptide fragment that targets a unique motif encompassing the histidine of a specific HMGR homolog offers the opportunity to be highly specific towards that homolog. Similar to traditional drug design and development that targets the active site, the proof-of-concept of using allosteric regulation still requires further scrutiny into the effects of using allosteric modulators, but the basic science research presented here provides some of the foundations of these approaches.

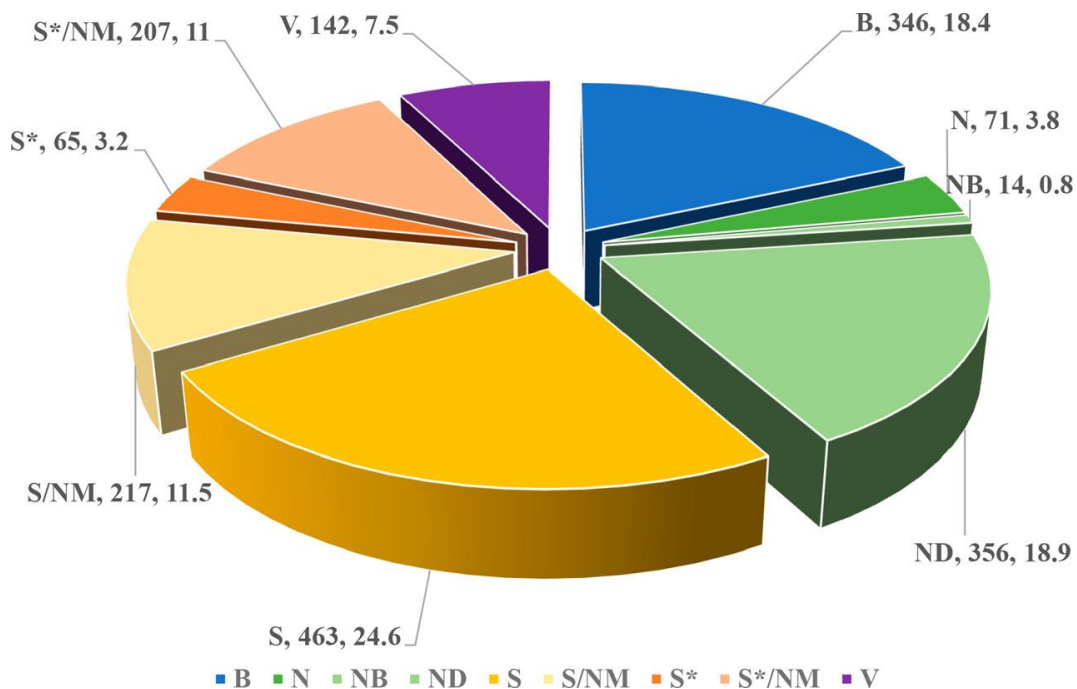
In this section, we discussed the importance of understanding the basic biochemistry of proteins in order to exploit them as drug targets. More specifically, by having intimate knowledge of the active site, allosteric sites, oligomerization motifs, and CTD dynamics, of HMGRs, one could design drug compounds that perturb activity as a means of inhibition. Greater understanding

of these endeavor's provides flexibility in drug design to be able to modify drug leads by balancing important drug efficacy parameters that eventually contribute towards its safety and effectiveness. The kinetic, biophysical, and crystallographic data presented in the earlier chapters were discussed in this context and proved valuable in expanding the preliminary work needed to exploit HMGR as a drug target.

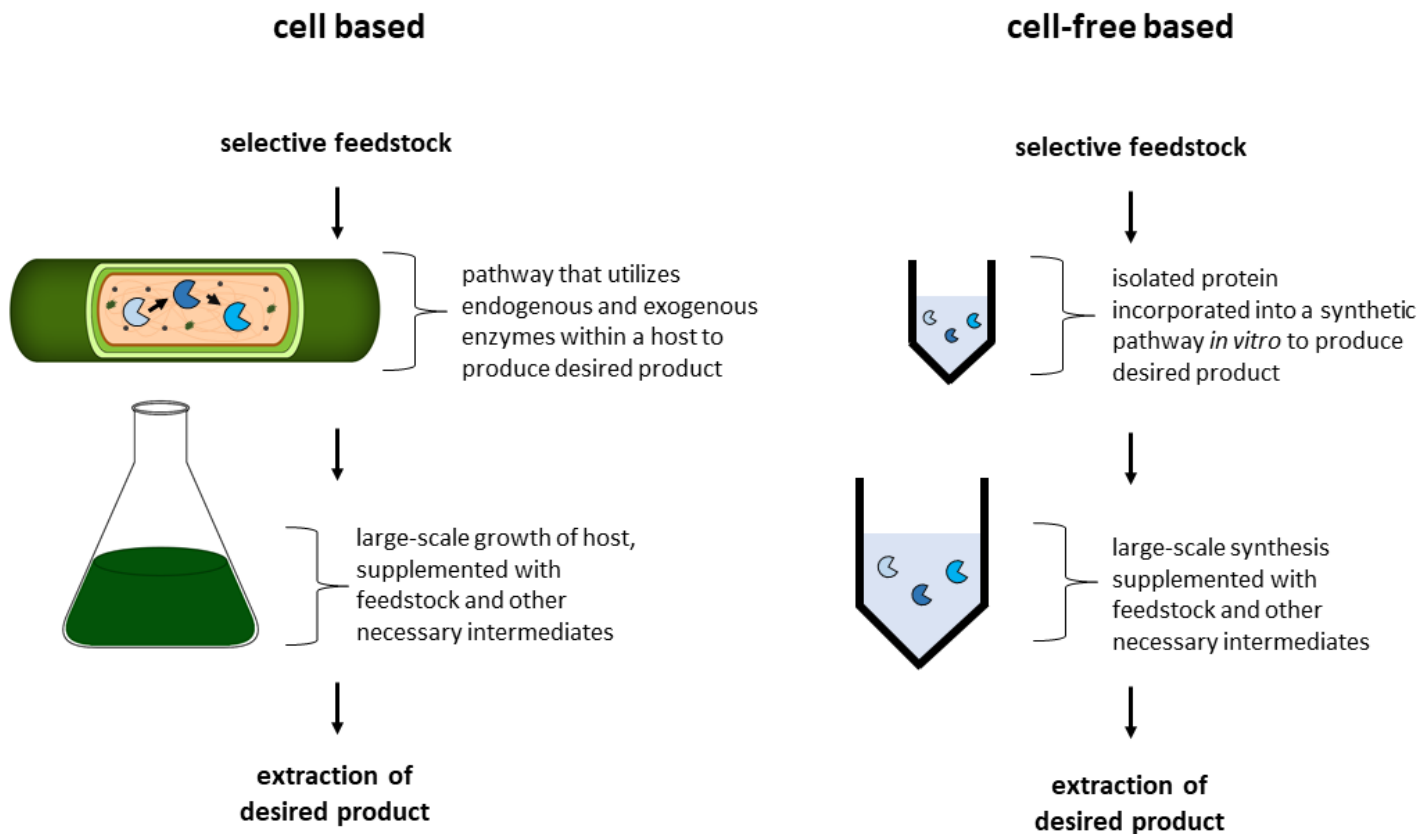
## **5.6. Conclusion**

In sum, this study adds to and extends our understanding of class II HMGR, an enzyme that is implicated across biology as well as in synthetic endeavors. This research has elucidated some of nature's terrific strategies in being able to generate unique structural features that give HMGR its ability to carry out its mechanism. It also offers biochemists the opportunity to harness these strategies for numerous benefits. Chapter Two highlighted the importance of biophysical characteristics such as oligomeric state in solution and overall three-dimensional dynamics. Chapter Three investigated the atomic-level detail of the mechanism and paid specific attention to the mobile C-terminal domain that directly aids in catalysis. Chapter Four focused on the structural determinants of cofactor specificity with protein engineering that utilized nature's design to rationally alter cofactor preferences. Despite significant strides being achieved, this study unearths further opportunities for inquiry and warrants ongoing investigation into this biologically important enzyme. Its direct applications are abundant, but, importantly, the advancement in our understanding of the way evolution has developed this fascinating protein continues to be scientifically fruitful and is valuable as we pursue deeper understanding of nature's grand design.

## 5.7. Tables and Figures

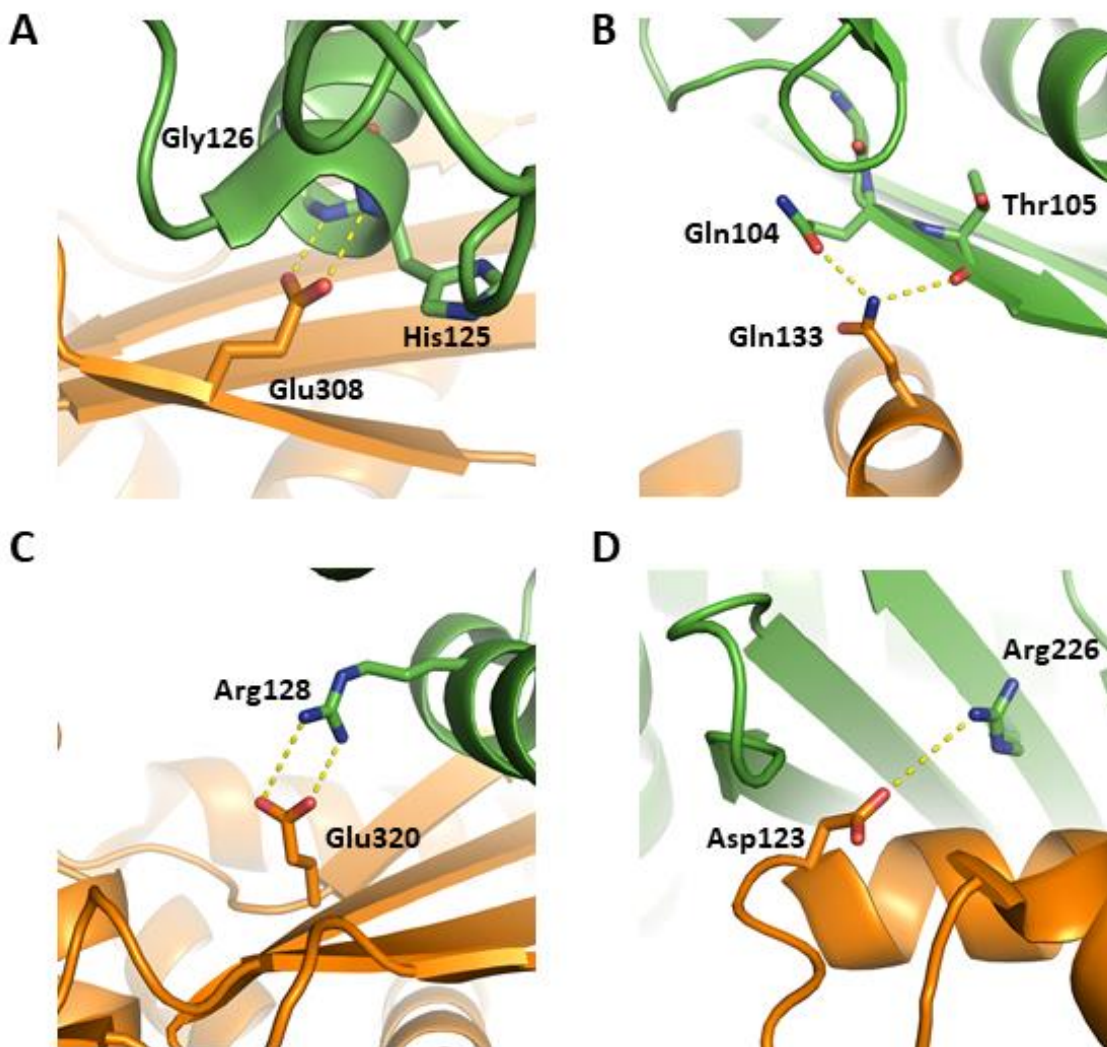


**Figure 1. Small-molecule drugs approved from 01/1981 to 09/2019.** Newman and Cragg highlight the small molecules that were approved based on their source in their pie chart in this figure. *The categories as outlined by the authors are as follows:* B = Biological, N= Natural product, NB= Natural product "Botanical", ND = Derived from a natural product, S = Totally synthetic drug, S\* = Made by total synthesis, but the pharmacophore is/was from a natural product, V = vaccine. (Newman and Cragg, *J. Nat. Prod.* 2020, 83, 3, 770–803)

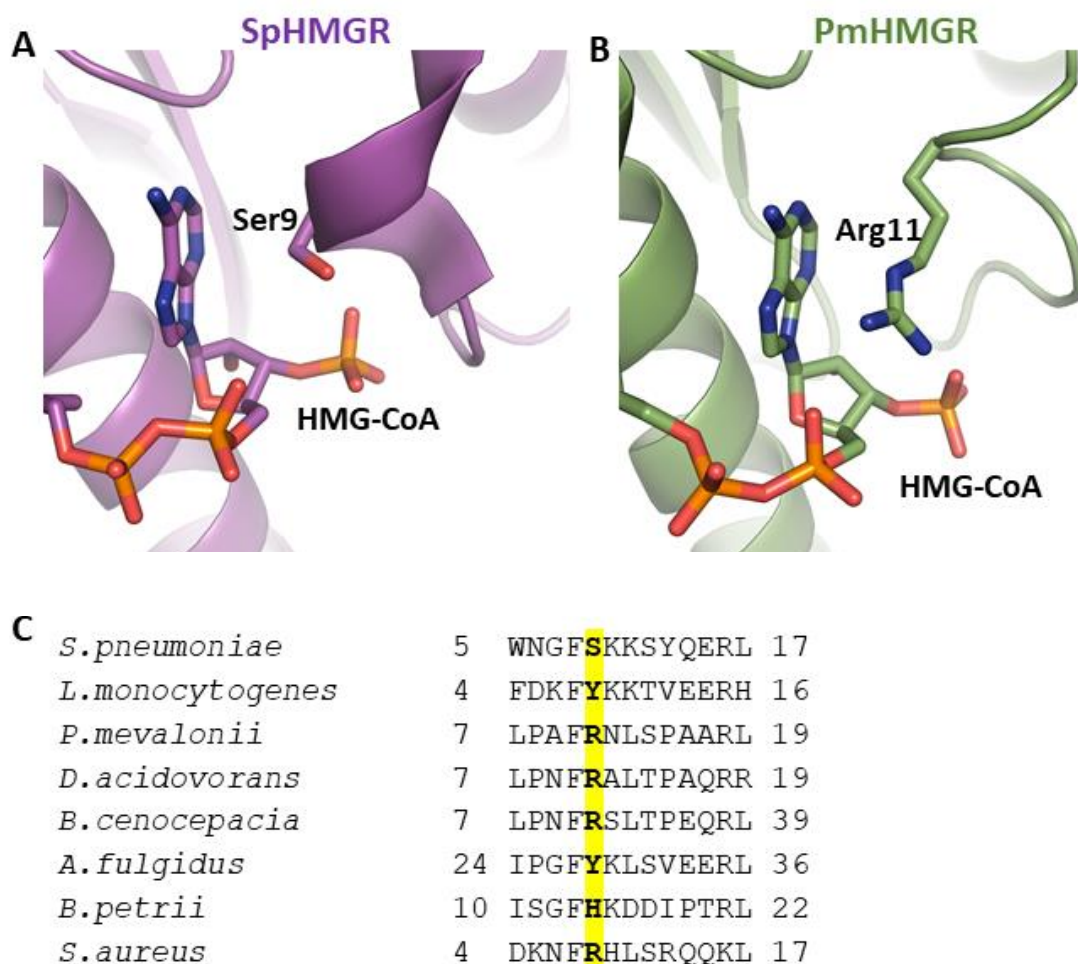


**Figure 2. Cell based versus cell-free based metabolic engineering.** Cell based metabolic engineering (left), utilizes endogenous and exogenous enzymes of a pathway within a host, to convert selective feedstock to a desired product. The host cell can be grown in large amounts and may require supplementation with reaction intermediates such that large-scale production of the desired product is obtained and subsequently extracted. Cell-free based metabolic engineering (right) is done *in vitro* where isolated proteins are incorporated into a synthetic pathway, that utilizes the selective feedstock as a substrate, to create the desired product. Large-scale synthesis can be obtained and may require supplementation of necessary metabolites to create large amounts of the desired product that can subsequently be extracted.

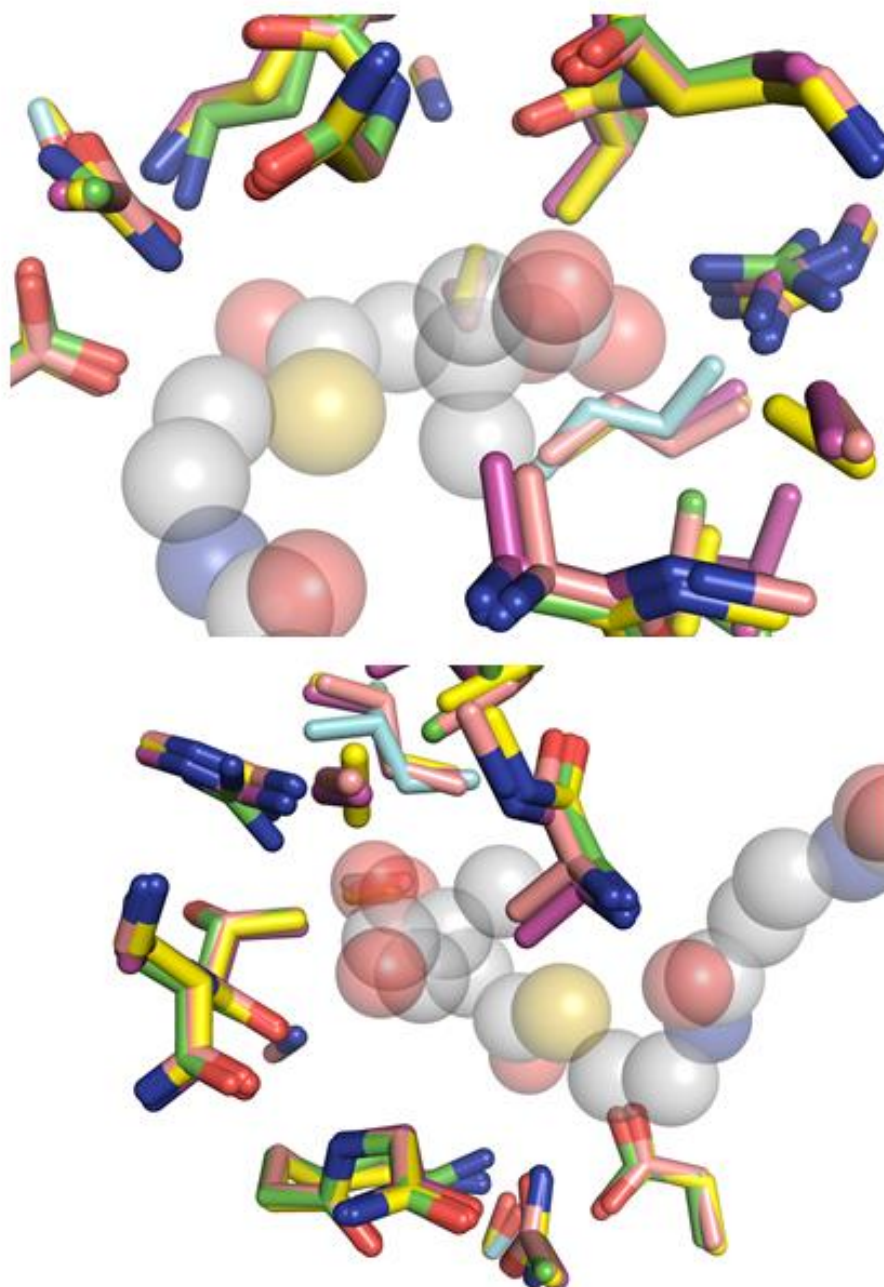




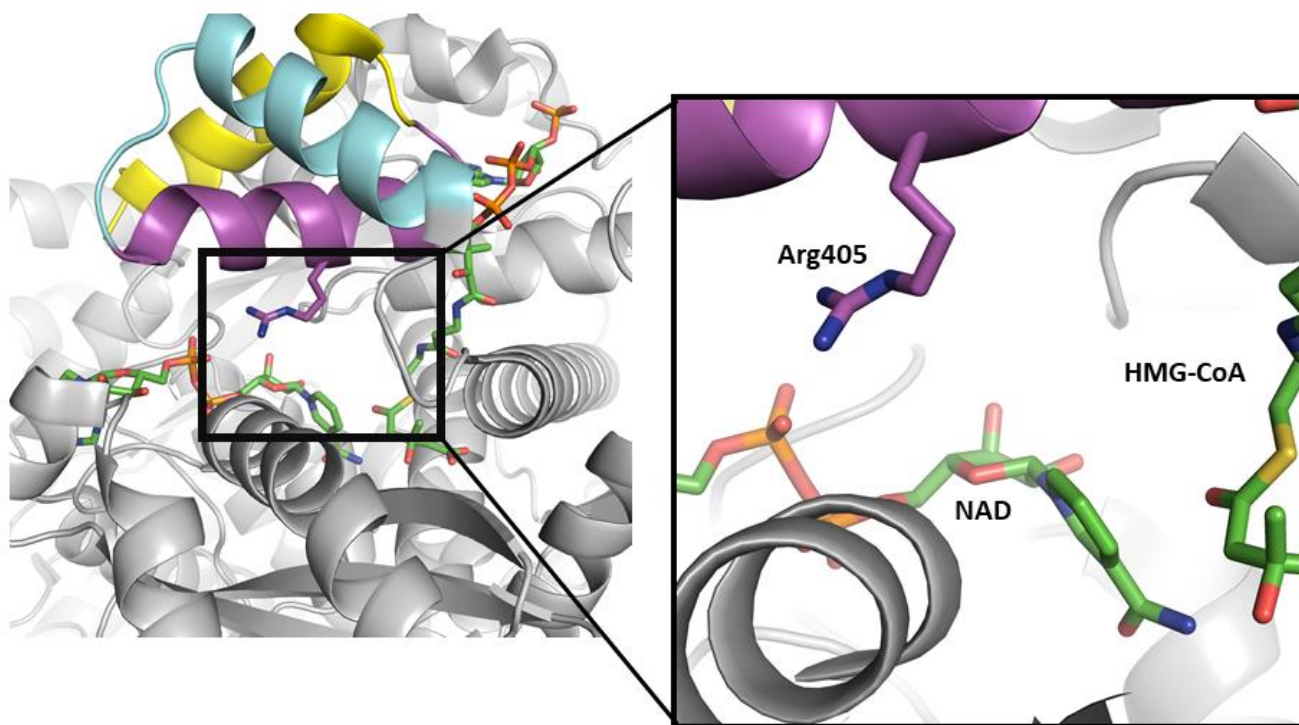
**Figure 3. Hexameric interface interactions.** Various interactions between adjacent monomers at the hexameric interface of DaHMGR. These interfaces appear to stabilize the hexameric formation and are candidates for site-directed mutagenesis when trying to modulate the formation of the hexamer. One monomer is colored orange, and the adjacent monomer is colored green. When depicted, residues and backbone are shown as sticks, with the following coloring: O = red, N = blue and carbons colored the same as its respective monomer. A) Glu320 makes backbone interactions with His125 and Gly126. B) Gln133 makes interactions with the side chain of Gln104 and with the backbone of Thr105. C) Glu320 makes electrostatic interactions with Arg128. D) A potential side chain interaction between Asp123 and Arg226. A) and B) are being investigated in the Kung Lab (Ashley Yang and Yan Kung, unpublished). C) and D) have been proposed by Peacock et al.



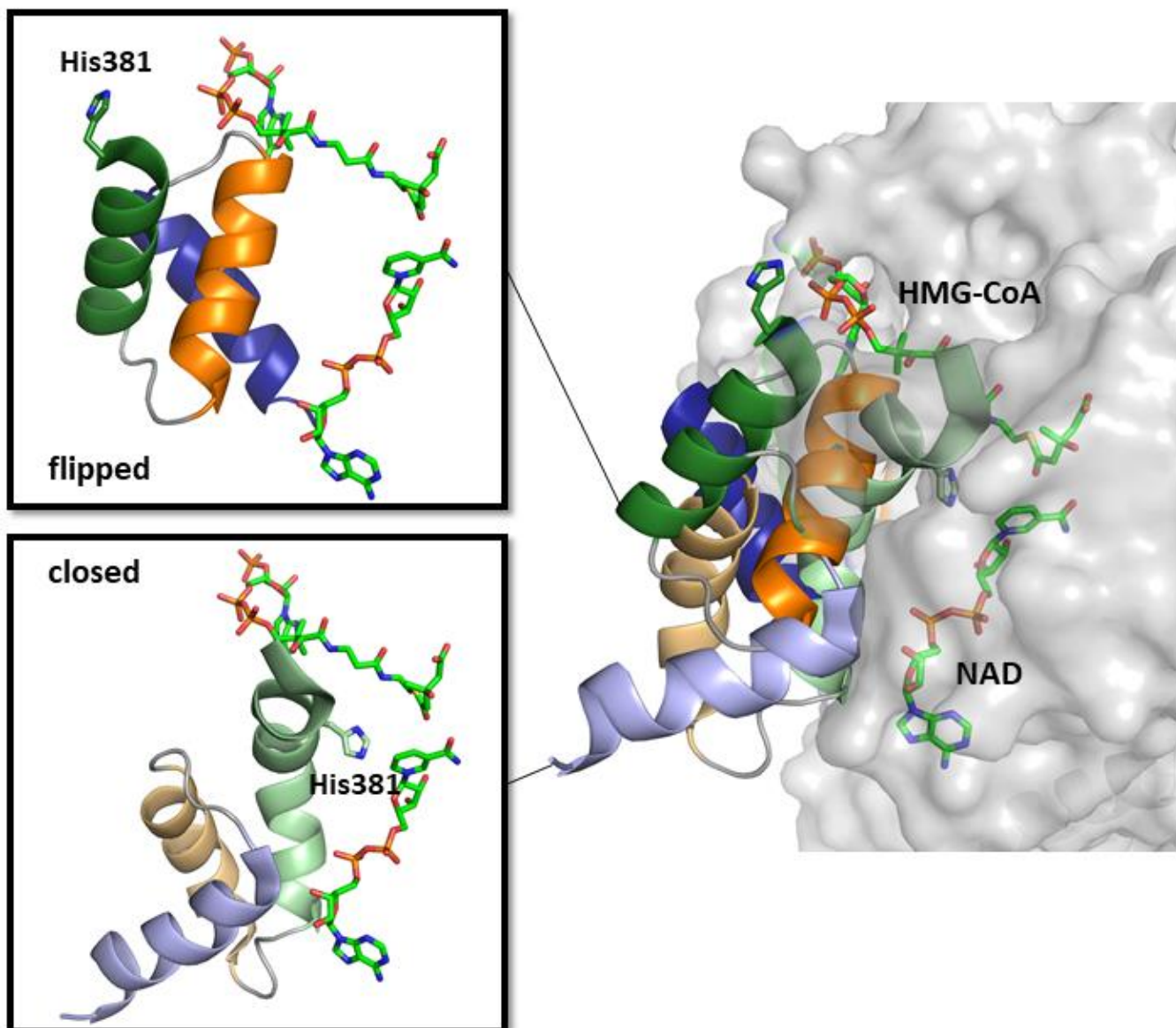
**Figure 4. Variance in residues that appear to interact with the solvent exposed adenine moiety of HMG-CoA in various HMGR homologs.** A) and B) show residues flanking the adenine moiety of HMG-CoA. A) HMG-CoA-bound SpHMGR (PDB 5WPK) shown in cartoon representation in purple with Ser9 and HMG-CoA shown as sticks. B) HMG-CoA-bound PmHMGR (PDB 1QAX) shown in cartoon representation in green with Arg11 and HMG-CoA shown as sticks. Sticks are colored as follows: O = red, N = blue, C = colored according to the cartoon representation. C) A sequence alignment at the N-terminal region of various HMGR homologs, with the highlighted residue being the one that appears to flank the adenine moiety of HMG-CoA. The residue that appears to interact with the adenine appear to be either aromatic (tyrosine or histidine), arginine, or serine.



**Figure 5. Overlay of the active site of several class II HMGR homologs.** Two different views of an overlay showing the active site of HMGR of from various class II HMGRs. Side chains of residues in the active site are shown as sticks with the following coloring: PmHMGR in green (PDB 1QAX), SpHMGR in magenta (PDB 5WPJ), BcHMGR in yellow (PDB 6P7K), and DaHMGR in light pink (6EEU). HMG-CoA is shown as spheres and are colored as follows: C = gray, O = red, N = blue, S = yellow, P = orange.



**Figure 6. Arg405 introduced near the active site as a result of the flipped conformation found in citrate- and NAD-bound DaHMGR.** Citrate- and NAD-bound DaHMGR shown in cartoon representation colored gray, with the first, second and third helices of the CTD in the flipped conformation colored cyan, magenta, and yellow respectively. Arg405, which is part of the second helix, is now in the vicinity of the active site, closer to the cofactor. HMG-CoA is modelled in from HMG-CoA-bound PmHMGR (PDB 1QAX). Worth noting is that due to the large sidechain of Arg, this residue can be exploited as a drug target beyond its specific positioning above. Ligand sticks are colored as follows: C = green, O = red, P = orange, N = blue, S = yellow.



**Figure 7. Flipped versus closed conformation of the CTD.** On the right, is an overlay of the flipped and the closed conformations. The CTD of citrate- and NAD-bound DaHMGR (PDB 6DIO) in the flipped conformation is shown in cartoon representation with helices 1, 2, and 3 colored in dark green, dark orange, and dark blue respectively. The CTD of HMG-CoA- and NAD-bound PmHMGR (PDB 1QAX) in the closed conformation is shown in cartoon representation with helices 1, 2 and 3 colored light green, light orange, and light blue respectively. The rest of the protein of DaHMGR and PmHMGR is depicted in gray surface representation. NAD and HMG-CoA from HMG-CoA- and NAD-bound PmHMGR (PDB 1QAX) is modeled in. The catalytical histidines, His381, of each CTD is shown in sticks. On the left, is a close-up view of the flipped and closed conformations, highlighting the location of His381 in each CTD. In the flipped conformation, His381 is solvent exposed and away from the active site. In the closed conformation, His381 is buried and directed towards the active site. Sticks are colored as follows: C = green, N = blue, O = red, P = orange, and S = yellow.

Table 1.  $K_m$  values for various class II HMGRs shown.

<b>HMGR homolog</b>	<b><math>K_M</math> HMG-CoA (<math>\mu</math>M)</b>	<b><math>K_M</math> NADH (<math>\mu</math>M)</b>
<b>DaHMGR</b>	7.34	28.5
<b>LmHMGR</b>	19.8	150
<b>EfHMGR</b>	20	ND
<b>SaHMGR</b>	40	100
<b>PmHMGR</b>	50	80
<b>SpHMGR</b>	75.86	153
<b>AfHMGR</b>	175	160

## 5.8. References

- [1] D. A. Dias, S. Urban, and U. Roessner, “A Historical Overview of Natural Products in Drug Discovery,” *Metabolites*, vol. 2, no. 2, pp. 303–336, Apr. 2012, doi: 10.3390/metabo2020303.
- [2] D. J. Newman and G. M. Cragg, “Natural Products as Sources of New Drugs from 1981 to 2014,” *Journal of Natural Products*, vol. 79, no. 3, pp. 629–661, 2016, doi: 10.1021/acs.jnatprod.5b01055.
- [3] D. J. Newman and G. M. Cragg, “Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019,” *J. Nat. Prod.*, vol. 83, no. 3, pp. 770–803, Mar. 2020, doi: 10.1021/acs.jnatprod.9b01285.
- [4] J. D. Keasling, “Manufacturing Molecules Through Metabolic Engineering,” *Science*, vol. 330, no. 6009, p. 1355, Dec. 2010, doi: 10.1126/science.1193990.
- [5] J.-W. Lee and C. T. Trinh, “Towards renewable flavors, fragrances, and beyond,” *Current Opinion in Biotechnology*, vol. 61, pp. 168–180, Feb. 2020, doi: 10.1016/j.copbio.2019.12.017.
- [6] Y. Boucher, H. Huber, S. L’Haridon, K. O. Stetter, and W. F. Doolittle, “Bacterial Origin for the Isoprenoid Biosynthesis Enzyme HMG-CoA Reductase of the Archaeal Orders Thermoplasmatales and Archaeoglobales,” *Molecular Biology and Evolution*, vol. 18, no. 7, pp. 1378–1388, Jul. 2001, doi: 10.1093/oxfordjournals.molbev.a003922.
- [7] D. S. Scher and V. W. Rodwell, “3-Hydroxy-3-methylglutaryl coenzyme A lyase from *Pseudomonas mevalonii*,” *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, vol. 1003, no. 3, pp. 321–326, Jun. 1989, doi: 10.1016/0005-2760(89)90239-7.
- [8] J. A. Friesen, C. M. Lawrence, C. V. Stauffacher, and V. W. Rodwell, “Structural Determinants of Nucleotide Coenzyme Specificity in the Distinctive Dinucleotide Binding Fold of HMG-CoA Reductase from *Pseudomonas mevalonii* †,” *Biochemistry*, vol. 35, no. 37, pp. 11945–11950, Jan. 1996, doi: 10.1021/bi9609937.
- [9] J. Zhang, N. Sonnenschein, T. P. B. Pihl, K. R. Pedersen, M. K. Jensen, and J. D. Keasling, “Engineering an NADPH/NADP<sup>+</sup> Redox Biosensor in Yeast,” *ACS Synth. Biol.*, vol. 5, no. 12, pp. 1546–1556, Dec. 2016, doi: 10.1021/acssynbio.6b00135.
- [10] E. I. Wilding *et al.*, “Identification, Evolution, and Essentiality of the Mevalonate Pathway for Isopentenyl Diphosphate Biosynthesis in Gram-Positive Cocci,” *J Bacteriol*, vol. 182, no. 15, pp. 4319–4327, Aug. 2000, doi: 10.1128/JB.182.15.4319-4327.2000.
- [11] B. R. Miller and Y. Kung, “Structural Features and Domain Movements Controlling Substrate Binding and Cofactor Specificity in Class II HMG-CoA Reductase,” *Biochemistry*, vol. 57, no. 5, pp. 654–662, Feb. 2018, doi: 10.1021/acs.biochem.7b00999.

- [12] T. C. Jordan-Starck and V. W. Rodwell, "Pseudomonas mevalonii 3-hydroxy-3-methylglutaryl-CoA reductase. Characterization and chemical modification," *J Biol Chem*, vol. 264, no. 30, pp. 17913–17918, Oct. 1989.
- [13] E. I. Wilding *et al.*, "Essentiality, Expression, and Characterization of the Class II 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase of *Staphylococcus aureus*," *J Bacteriol*, vol. 182, no. 18, pp. 5147–5152, Sep. 2000, doi: 10.1128/JB.182.18.5147-5152.2000.
- [14] J. Nielsen and J. D. Keasling, "Engineering Cellular Metabolism," *Cell*, vol. 164, no. 6, pp. 1185–1197, Mar. 2016, doi: 10.1016/j.cell.2016.02.004.
- [15] L. B. Pickens, Y. Tang, and Y.-H. Chooi, "Metabolic Engineering for the Production of Natural Products," *Annu. Rev. Chem. Biomol. Eng.*, vol. 2, no. 1, pp. 211–236, Jul. 2011, doi: 10.1146/annurev-chembioeng-061010-114209.
- [16] H. J. Lim and D.-M. Kim, "Cell-Free Metabolic Engineering: Recent Developments and Future Prospects," *MPs*, vol. 2, no. 2, p. 33, Apr. 2019, doi: 10.3390/mps2020033.
- [17] R. W. Haushalter *et al.*, "Production of anteiso-branched fatty acids in *Escherichia coli*; next generation biofuels with improved cold-flow properties," *Metabolic Engineering*, vol. 26, pp. 111–118, Nov. 2014, doi: 10.1016/j.ymben.2014.09.002.
- [18] S. M. Ma *et al.*, "Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases," *Metabolic Engineering*, vol. 13, no. 5, pp. 588–597, Sep. 2011, doi: 10.1016/j.ymben.2011.07.001.
- [19] R. B. Peacock *et al.*, "Structural and Functional Characterization of Dynamic Oligomerization in *Burkholderia cenocepacia* HMG-CoA Reductase," *Biochemistry*, vol. 58, no. 38, pp. 3960–3970, Sep. 2019, doi: 10.1021/acs.biochem.9b00494.
- [20] G. L. Patrick, *An introduction to medicinal chemistry*, Fifth edition. Oxford: Oxford University Press, 2013.
- [21] I. Gashaw, P. Ellinghaus, A. Sommer, and K. Asadullah, "What makes a good drug target?," *Drug Discovery Today*, vol. 16, no. 23–24, pp. 1037–1043, Dec. 2011, doi: 10.1016/j.drudis.2011.09.007.
- [22] J. Drews, "Basic science and pharmaceutical innovation," *Nat Biotechnol*, vol. 17, no. 5, pp. 406–406, May 1999, doi: 10.1038/8531.
- [23] T. A. Taylor and C. G. Unakal, "Staphylococcus Aureus," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2021. Accessed: Nov. 10, 2021. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK441868/>
- [24] M. Hedl and V. W. Rodwell, "Inhibition of the Class II HMG-CoA reductase of *Pseudomonas mevalonii*," *Protein Sci.*, vol. 13, no. 6, pp. 1693–1697, Jun. 2004, doi: 10.1110/ps.03597504.



- [25] E. Istvan, "Bacterial and mammalian HMG-CoA reductases: related enzymes with distinct architectures," *Current Opinion in Structural Biology*, vol. 11, no. 6, pp. 746–751, Dec. 2001, doi: 10.1016/S0959-440X(01)00276-7.
- [26] Grace M. Fimognari and Victor W. Rodwell, "Substrate-competitive inhibition of bacterial mevalonate: nicotinamide-adenine dinucleotide oxidoreductase (acylating CoA)," *Biochemistry*, vol. 4, no. 10, pp. 2086–2090, 1965.
- [27] R. M. Karl, M. Hupperich, A. Thomer, and H. Eggerer, "Inhibition of 3-Hydroxy-3-Methylglutaryl-CoA Reductase and Citrate Synthase by Sulfoxides and Sulfones of Substrate-analogue CoA-thioether Derivatives," *Eur J Biochem*, vol. 227, no. 1–2, pp. 292–295, Jan. 1995, doi: 10.1111/j.1432-1033.1995.tb20387.x.
- [28] L. Tabernero, D. A. Bochar, V. W. Rodwell, and C. V. Stauffacher, "Substrate-induced closure of the flap domain in the ternary complex structures provides insights into the mechanism of catalysis by 3-hydroxy-3-methylglutaryl-CoA reductase," *Proceedings of the National Academy of Sciences*, vol. 96, no. 13, pp. 7167–7171, Jun. 1999, doi: 10.1073/pnas.96.13.7167.
- [29] A. K. Grover, "Use of Allosteric Targets in the Discovery of Safer Drugs," *Med Princ Pract*, vol. 22, no. 5, pp. 418–426, 2013, doi: 10.1159/000350417.
- [30] C. J. Wenthur, P. R. Gentry, T. P. Mathews, and C. W. Lindsley, "Drugs for Allosteric Sites on Receptors," *Annu. Rev. Pharmacol. Toxicol.*, vol. 54, no. 1, pp. 165–184, Jan. 2014, doi: 10.1146/annurev-pharmtox-010611-134525.