Understanding Reductive Enzymes in the Erythritol Biosynthesis Pathway:
"The Secrets of Sweet Blood"

A Thesis

Presented to the Biology & Society Program

of Cornell University

for Honors Thesis Consideration

by

Peyton Lee Carpen
May 3rd, 2023

Table of Contents

Abstract:

Erythritol is a sugar alcohol and non-nutritive sweetener that the body produces from glucose. Studies have shown a correlation between cardiometabolic dysfunction and increased erythritol production. It is hypothesized that elevated erythritol is a predictive biomarker of cardiometabolic diseases. The overall objectives of this project are to uncover the characteristics of enzymes catalyzing the erythritol biosynthesis pathway through wet lab and computational experiments and to understand how this pathway is causally related to development of cardiometabolic dysfunction. Two human enzymes have been identified as catalyzing the final step in the conversion of glucose to erythritol: *ADH1* (alcohol dehydrogenase) & *SORD* (sorbitol dehydrogenase.)[1] In order to interrogate how these enzymes relate to whole-body cardiometabolic dysfunction, a mouse model is required. The recombinant murine (mouse) ADH1 and SORD enzymes were cloned and purified to discern whether the murine enzymes also catalyze the final step in erythritol production. The enzymatic activity found was comparable between species. However, when mouse models were generated lacking either *Adh1* or *Sord*, there was no robust decrease in erythritol production in mice.[2] These results suggest that enzymes other than ADH1 and SORD catalyze this same activity in mice. Therefore, a computational search was conducted to identify alternative enzymes with structural similarity to ADH1 and SORD that may catalyze this reaction in mice and in humans. ADH5 was one identified as a potential alternate enzyme of interest for future study.

[1] Schlicker L, Szebenyi DME, Ortiz SR, Heinz A, Hiller K, Field MS. Unexpected roles for ADH1 and SORD in catalyzing the final step of erythritol biosynthesis. J Biol Chem. 2019 Nov 1;294(44):16095-16108. doi: 10.1074/jbc.RA119.009049. Epub 2019 Sep 11. PMID: 31511322; PMCID: PMC6827307.
(Note: Schlicker, et al. 2019 citation will be used)

[2] Elevated plasma and urinary erythritol is a biomarker of excess simple carbohydrate intake in mice
Semira R. Ortiz, Martha S. Field bioRxiv 2022.12.04.519026; doi: https://doi.org/10.1101/2022.12.04.519026

Literature Review:

Erythritol is a common sugar alcohol and food additive.[3] Sugar alcohols, or [4]polyols, are often used as an alternative to sugar. Erythritol is a non-nutritive sweetener and does not spike insulin levels as traditional table sugar does.[5] Erythritol is a 4-carbon polyol that can be found in a myriad of foods, both processed and unprocessed. It has the highest concentrations in processed, low-calorie foods such as 100-calorie snack packs and diet beverages. It can also be found in lower amounts in non-processed or fermented foods including various fruits, wines, and cheeses. It is safe for human consumption, and its use in processed food products and drinks has exploded over the past 10 years. Intriguingly, researchers have discovered that erythritol is also endogenously produced by the human body.[6] Previously, endogenous erythritol production was only thought to occur in various yeast, bacterial and plant species.[7] "Endogenous production" in humans means that erythritol is produced by the body, rather than a result of exogenous, or external dietary, consumption. One of first papers to identify erythritol as being produced by the body also identified elevated plasma erythritol levels as associated with increased risk of elevated glycemia and increased abdominal adiposity.[89] These conditions can impair the insulin

[3] Curtis L. Fritz, Duc J. Vugia, Chapter 14 - Brucellosis, Editor(s): J. Glenn Morris, Duc J. Vugia, Foodborne Infections and Intoxications (Fifth Edition), Academic Press, 2021,Pages 253-264,ISBN 9780128195192, https://doi.org/10.1016/B978-0-12-819519-2.00016-5.
[4] Karl F. Tiefenbacher, in The Technology of Wafers and Waffles II, 2019
[5] *Sugar Alcohol - Yale New Haven Hospital*, https://www.ynhh.org/services/nutrition/sugar-alcohol.
[6] Hootman KC, Trezzi JP, Kraemer L, Burwell LS, Dong X, Guertin KA, Jaeger C, Stover PJ, Hiller K, Cassano PA. Erythritol is a pentose-phosphate pathway metabolite and associated with adiposity gain in young adults. Proc Natl. (Note: Hootman, et al., 2017 will be used)
[7] Dorota A. Rzechonek, Adam Dobrowolski, Waldemar Rymowicz & Aleksandra M. Mirończuk (2018) Recent advances in biological production of erythritol, Critical Reviews in Biotechnology, 38:4, 620-633, DOI: 10.1080/07388551.2017.1380598=
[8] Ortiz SR, Field MS. Mammalian metabolism of erythritol: a predictive biomarker of metabolic dysfunction. Curr Opin Clin Nutr Metab Care. 2020
[9] Hootman, et al., 2017

response system and increase the likelihood of 'heart attacks, strokes, high blood pressure, cancer, diabetes, osteoarthritis, fatty liver and depression.'[10]

Rebholtz, et al.,[11] analyzed the plasma metabolome from a 20-year longitudinal study from a relatively large, diverse human cohort. This study identified seven compounds that when elevated in plasma had a strong association with higher risks of diabetes. One of the identified compounds was erythritol. Erythritol also has been identified in other studies as a molecule of interest in predicting risk for pathologies including coronary heart disease (CHD) [12] and central adiposity gain.[13] Wang, et al., demonstrated that that plasma erythritol and its metabolites were correlated with CHD risk over a 30-year period.[14] CHD causes heart attack and heart failure and often results in increased morbidity and mortality. Hootman, et. al., found in a cohort of freshman college students that those who showed central adiposity gain over the course of their freshman year exhibited higher levels of plasma erythritol at baseline, suggesting that elevated erythritol is a predictive biomarker of central adiposity gain. Central adiposity gain is associated with cardiometabolic diseases and their severity[15], emphasizing the importance of elevated plasma erythritol as a predictive biomarker.

The regulatory mechanisms governing the endogenous production of erythritol from glucose are still not fully understood or characterized. Similarly, whether elevated plasma erythritol levels play a causal role in the development of these diseases is also unknown. These

[10] Harvard Health Publishing. "Abdominal Obesity and Your Health - Harvard Health." *Harvard Health*, Harvard Health, 20 Jan. 2017, www.health.harvard.edu/staying-healthy/abdominal-obesity-and-your-health.
[11] Rebholz, et al., 2018
[12] Wang, Zhe, and Bing Yu. "Response Letter Regarding Article, 'Metabolomic Pattern Predicts Incident Coronary Heart Disease.'" *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 39, no. 8, 2019, https://doi.org/10.1161/atvbaha.119.313013.
[13] Hamed Kianmehr, et al. JAMA Netw Open. 2022;5(4):e227705. doi:10.1001/jamanetworkopen.2022.7705
[14] Wang, Zhe, and Bing Yu. "Response Letter Regarding Article, 'Metabolomic Pattern Predicts Incident Coronary Heart Disease.'" *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 39, no. 8, 2019,
[15] Hamed Kianmehr, et al. JAMA Netw Open. 2022;5(4):e227705. doi:10.1001/jamanetworkopen.2022.7705

important questions cannot be answered without a more thorough understanding of the limiting substrates and enzymes in this pathway. The purpose of this study is to understand one piece of the erythritol biochemical puzzle.

It is also important to note public health implications of monitoring erythritol and of the greater body of literature focused on determining the properties of erythritol and characterizing its biosynthetic pathway. Erythritol has the potential to be a predictive biomarker for cardiometabolic diseases. Biomarkers are defined as any 'cellular, biochemical of molecular' measurable change in 'tissues, cells, or fluids.' Measuring these biomarkers can help establish when an individual is at risk for an illness or is already suffering from an abnormal pathology.[16] Advances in technology and increases in biomarker funding has helped researchers make great leaps in the field, and erythritol has shown great promise in being the next big biomarker in nutrition research.

Specifically, blood plasma erythritol analysis could help determine healthy baseline levels for different patient populations as well as monitor circulating serum erythritol levels over time to gauge risk levels for metabolic dysfunction before disease onset. This is important because existing biomarkers for cardiometabolic diseases and their progression appear only *after* the onset of the disease. Cardiometabolic diseases (i.e., heart disease & diabetes) are the top causes of death in the United States.[17] Cardiovascular diseases, like strokes and heart failure, cost the US $378.0 billion dollars in 2018.[18] Diabetes risk is also predicted by elevated plasma erythritol, and these currently affect over 130 million Americans or one-third of the US

[16] Mayeux, R. Biomarkers: Potential uses and limitations. *Neurotherapeutics* **1**, 182–188 (2004). https://doi.org/10.1602/neurorx.1.2.182
[17] Rebholz, et al., 2018
[18] Tsao, Connie W., et al. "Heart Disease and Stroke Statistics—2022 Update: A Report from the American Heart Association." *Circulation,* vol. 145, no. 8, 2022, https://doi.org/10.1161/cir.0000000000001052.

population. [19] New and innovative ways to prevent the onset of such deadly and costly diseases are desperately needed. Thus, a predictive biomarker such as erythritol could have clinical use in precision nutrition and preventive medicine. For example, in the future, measuring blood plasma erythritol could be just as common as glycated hemoglobin (A1C) test. The A1C test is a widely used blood test used to diagnose type 1 and type 2 diabetes, and diagnosis is determined by the blood glucose levels or AIC levels. [20] High glucose and A1C levels indicate endocrinological issues that must be treated with diabetes medication. However, this test is mainly used for diagnosis and treatment, rather than disease prevention. With a predictive biomarker such as erythritol, patients and medical practitioners will have additional time to identify those at risk and prevent these individuals from developing life-threatening illnesses. To reach that end goal, we must better understand the unique connection between elevated plasma erythritol and cardiometabolic dysfunction (Figure 1).

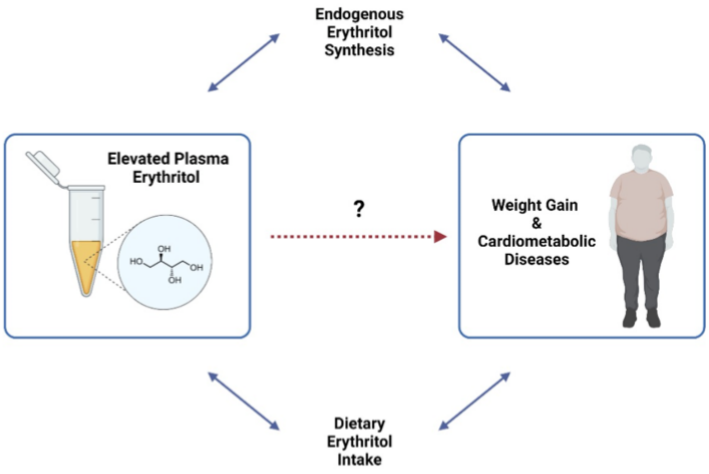
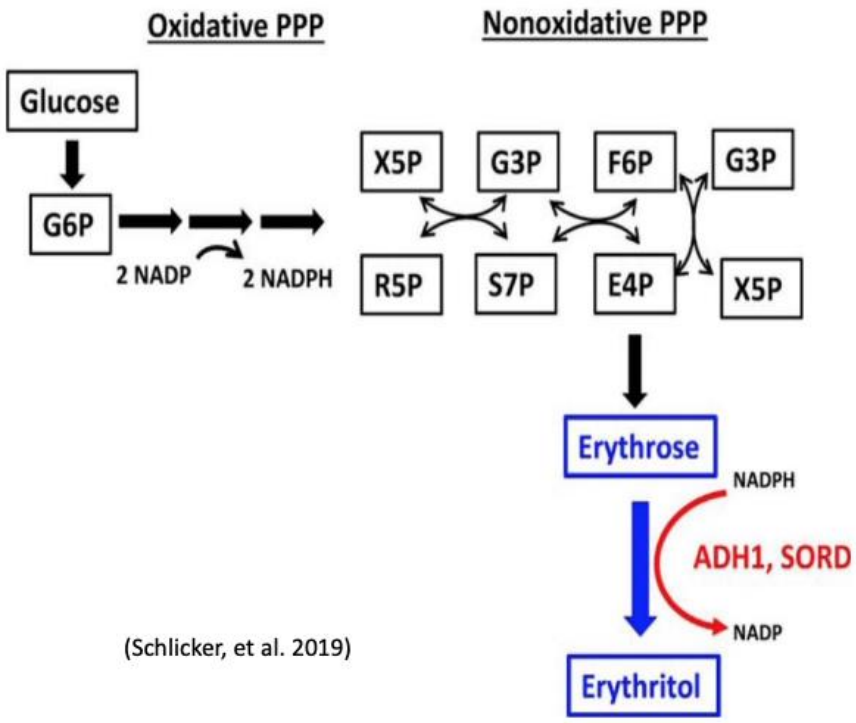


Figure 1: Diagram outlining the key scientific question addressed in this work

[19] Tsao, Connie W., et al. "Heart Disease and Stroke Statistics—2022 Update: A Report from the American Heart Association." *Circulation*, vol. 145, no. 8, 2022, https://doi.org/10.1161/cir.0000000000001052.
[20] "A1C Test." *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 1 Dec. 2022, https://www.mayoclinic.org/tests-procedures/a1c-test/about/pac-20384643.

As indicated above, the discovery that humans produce erythritol from glucose is relatively recent. In 2017 Hootman et. al. discovered that erythritol is synthesized using the pentose phosphate pathway (PPP) using stable isotope tracing experiments.[21] The PPP is a vital metabolic pathway as it produces reducing equivalents (in the form of NADPH) and other biosynthetic intermediates including ribose-5-phosphate for DNA synthesis.[22] However, the enzymes catalyzing the final step in this conversion of glucose to erythritol were unknown at that time. Subsequent studies isolated two enzymes from rabbit liver capable of catalyzing the last step in the conversion of glucose to erythritol: sorbitol dehydrogenase (SORD) and alcohol dehydrogenase 1 (ADH1).[23]



[21] Hootman, et al., 2017
[22] Garrett, H., Reginald, and Charles Grisham. Biochemistry. Boston: Twayne Publishers, 2008.
[23] Schlicker, et al., 2019

Figure 2: ***Conversion of glucose to erythritol via the pentose phosphate pathway.*** *Stable isotope tracers indicated that glucose was converted to erythritol through the pentose phosphate pathway in human cells These studies also demonstrated that the substrate for erythritol synthesis is erythrose, not erythrose-4-phosphate, and that the required cofactor in human cells is NADPH. In addition, this reaction is catalyzed by ADH1 and SORD. F6P, fructose 6-phosphate; G3P, glyceraldehyde 3-phosphate; G6P, glucose 6-phosphate; R5P, ribose 5-phosphate; S7P, sedoheptulose 7-phosphate; X5P, xylulose 5-phosphate." (Schlicker, 2019)* [24][25]

ADH1 participates in several metabolic pathways and disease pathologies, including alcohol metabolism, cancer development, liver disease, gout, asthma, fetal alcohol syndrome and CHD. [26] *ADH1* has several common genetic variants, and it is highly expressed in liver and stomach tissues. Two variants of interest are *ADH1B1* and *ADH1B2*, which had differing erythrose reduction activities. Specifically, *ADH1B2* had little to no reductive activity, whereas ADH1B1 was active and exhibited substrate inhibition, which is a phenomenon that occurs in around 25% of dehydrogenase enzymes.[27] Substrate inhibition is the formation of an unproductive enzyme-substrate complex after the simultaneous binding of two or more substrate molecules to the active site.[28] ADH1 variants have different reactivities in ethanol metabolism.

[24] Schlicker, et al., 2019
[25] Schlicker, et al., 2019
[26] Crabb DW, Matsumoto M, Chang D, You M. Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. Proc Nutr Soc. 2004 Feb;63(1):49-63. doi: 10.1079/pns2003327. PMID: 15099407.
[27] Kokkonen P, Beier A, Mazurenko S, Damborsky J, Bednar D, Prokop Z. Substrate inhibition by the blockage of product release and its control by tunnel engineering. RSC Chem Biol. 2021 Jan 11;2(2):645-655. doi: 10.1039/d0cb00171f. PMID: 34458806; PMCID: PMC8341658
[28] Kokkonen P, Beier A, Mazurenko S, Damborsky J, Bednar D, Prokop Z. Substrate inhibition by the blockage of product release and its control by tunnel engineering. RSC Chem Biol. 2021 Jan 11;2(2):645-655. doi: 10.1039/d0cb00171f. PMID: 34458806; PMCID: PMC8341658

Whether these variants also having meaningful differences in erythritol synthesis capacity remains to be determined.

SORD serves a key role in carbohydrate and sugar metabolism.[29] It is primarily used to interconvert sorbitol and fructose as a part of the polyol pathway. The polyol pathway converts glucose into sorbitol and then fructose, and this pathway has been implicated in the disease progression of late-stage diabetes.[30] SORD is highly expressed in liver and kidney tissues. *SORD* has fewer known human genetic variants than *Adh1*.[31] Although *SORD* does not have the same diversity of known variants as *ADH1*, there are two variants that result from alternate gene splicing.[32] It is also important to note that ADH1 and SORD enzymes are 'structurally and kinetically homologous.'[33]

Intriguingly, Ortiz, et al. conducted an additional study demonstrating that decreased *SORD* expression in human cells (A549 lung cancer cells) resulted in a significant decrease in erythritol production, suggesting that this enzyme may be rate-limiting for erythritol synthesis from glucose in humans. *ADH1* was not expressed at high enough levels in A549 cells to determine the role of this enzyme in erythritol production *in vivo.*[34] Given these findings, a next logical step in the project was to use mouse models lacking either SORD or ADH1 proteins to

[29] El-Kabbani O, Darmanin C, Chung R P. Sorbitol dehydrogenase: structure, function and ligand design. *Curr Med Chem*, 2004, 11(4):465–476.
[30] Srikanth, KK, and JA Orrick . "Biochemistry, Polyol or Sorbitol Pathways - NCBI Bookshelf." *National Library of Medicine*, 2022, https://www.ncbi.nlm.nih.gov/books/NBK576381/.
[31] Schlicker, et al., 2019
[32] "Uniprot ." *UniProt*, https://www.uniprot.org/uniprotkb/Q00796/entry.
[33] Schlicker, et al., 2019
[34] Ortiz SR, Heinz A, Hiller K, Field MS. Erythritol synthesis is elevated in response to oxidative stress and regulated by the non-oxidative pentose phosphate pathway in A549 cells. Front Nutr. 2022 Oct 6;9:953056. doi: 10.3389/fnut.2022.953056. PMID: 36276829; PMCID: PMC9582529.

determine the effect of loss of these enzymes on whole-body erythritol metabolism. Mouse models are very commonly used to study obesity and cardiometabolic dysfunction[35].

Therefore, an *in vitro* study (presented herein) was conducted to determine if the enzymatic activity of recombinant mouse ADH1 and SORD was comparable to the activity of the human recombinant enzymes. As presented below (i.e. results), both recombinant murine ADH1 and SORD exhibited the ability to catalyze the final step in the conversion of glucose to erythritol. However, when these studies were followed-up *in vivo*, mouse models lacking either ADH1 or SORD did not exhibit decreased endogenous erythritol production.[36] This finding suggests that mice (and possibly humans) express alternative enzymes with the ability to catalyze the final step in the conversion of glucose to erythritol. Another possibility is that ADH1 and SORD compensate for one another, leading to the question of pursuing a double KO of SORD and ADH1. This is unlikely because of recent findings showing that other mechanisms like nfr2 oxidation and TKT are implicated in erythritol synthesis.[37] Also, there is evidence that SORD KO in A549 cells decreases erythritol production, without the presence of ADH1[38], effectively creating a double KO in cell culture. This suggests that there is an alternate enzyme of interest, besides SORD and ADH1. Thus, it was necessary to identify homologous structures for the two enzymes of interest. Homology, or sequence and structural similarity[39], can be found by investigating alternate enzymes. A BLASTx , or a nucleotide search that uses existing protein

[35] Speakman, J., Hambly, C., Mitchell, S. and Król, E. (2007), Animal models of obesity. Obesity Reviews, 8: 55-61. https://doi.org/10.1111/j.1467-789X.2007.00319.x
[36] Elevated plasma and urinary erythritol is a biomarker of excess simple carbohydrate intake in mice Semira R. Ortiz, Martha S. Field. bioRxiv 2022.12.04.519026; doi: https://doi.org/10.1101/2022.12.04.519026
[37] Ortiz SR, Heinz A, Hiller K, Field MS. Erythritol synthesis is elevated in response to oxidative stress and regulated by the non-oxidative pentose phosphate pathway in A549 cells. Front Nutr. 2022 Oct 6;9:953056. doi: 10.3389/fnut.2022.953056. PMID: 36276829; PMCID: PMC9582529.
[38] Schlicker, et al., 2019
[39] Koonin, Eugene V. (2005). Orthologs, Paralogs, and Evolutionary Genomics 1. https://doi.org/10.1146/annurev.genet.39.073003.114725

data bases to find sequence similarity[40] was preformed briefly and it yielded no alternate enzymes of interest. Thus, an additional computational structural similarity search for homology was conducted rather than a sequence similarity search. It was found that there was one potential alternate enzyme of interest for ADH1 and none for SORD. The identified alternate enzyme of interest for ADH1 was ADH5.

[40] Library, UC Berkeley. "Library Guides: NCBI Bioinformatics Resources: An Introduction: NCBI Homepage." *NCBI Homepage - NCBI Bioinformatics Resources: An Introduction - Library Guides at UC Berkeley*, https://guides.lib.berkeley.edu/ncbi.

Methodology:

Part 1: Compare the Enzymatic Activity of ADH1 & SORD in Human and Murine Models

Background:

Murine *Adh1* and *Sord* genes were cloned into pet28a(+) vectors. Cloning is a common molecular biology process that allows researchers to study proteins and enzymes of interest. Restriction endonucleases are used to 'cut' out a region of a circular bacterial genome, also known as a plasmid.[41] The pet 28+ vector has a recombinant gene of interest inserted, when bacteria is transfected with this vector it will reproduce and produce the protein of interest. DNA ligases 'glue' everything back together, giving a circular plasmid with the new gene of interest inserted.[42] In this case, there were two separate plasmids with the inserted *Adh1* and *Sord* genes, respectively. They both contain a 6-His tag, which allows the researcher to recover and extract the protein of interest in a large bacterial cell colony.[43] These bacterial plasmids were grown in a nutrient-rich media (Luria Broth, or LB) used to culture bacterial colonies.[44] To ensure successful cloning, Sanger sequencing, a process that verifies that the gene cloned in the pet28a+ vector (*Adh1* or *Sord*) is in frame with the 6XHis tag so that it can be successfully expressed and purified. Once this is achieved, bacteria are grown at large scale and then a purification step must occur to recover the cloned enzyme from the bacterial lysate. The enzymes, ADH1 and SORD were purified using Co^{2+} resin (beads), which bind to the "6xHis" tag on the recombinant

[41] Trigiano, Robert N., and Dennis J. Gray. "Chapter 14 - Techniques for Nucleic Acid Engineering: The Foundation of Gene Manipulation." *Plant Development and Biotechnology*, CRC, Boca Raton, FL, 2004.
[42] Watson JD (2007). *Recombinant DNA: genes and genomes: a short course*. San Francisco: W.H. Freeman. ISBN 978-0-7167-2866-5.
[43] Darain, F., Ban, C., & Shim, Y.-B. Development of a new and simple method for the detection of histidine tagged proteins. Biosensors and Bioelectronics. 2004:20(4);857–863. doi:10.1016/j.bios.2004.03.028.
[44] Addgene. "Addgene: Protocol - How to Inoculate a Bacterial Culture." *Addgene.org*, 2019, www.addgene.org/protocols/inoculate-bacterial-culture/.

enzyme. These beads are placed in a large vertical column. The bacterial lysate and enzyme mixture is slowly dripped into the column and run through the Co^{2+} resin. After a few hours, only the purified enzyme remains bound to the beads. The protein is eluted from the beads with a specialized wash buffer. If enough yield is obtained, this concentrated enzyme solution goes through one additional dialysis step to ensure that there are no additional small molecule contaminants like salts that could hinder future analysis.

Enzyme kinetic analysis is essential to both showing that there is a conversion of erythrose to erythritol via ADH1 and SORD enzymes and allowing for comparisons between murine recombinant enzymatic activity and human recombinant enzymatic activity. A cuvette filled with the purified enzyme (either ADH1 or SORD), an NADPH buffer and varying concentrations of erythrose was placed into the spectrophotometer. A spectrophotometer is used to analyze the decrease in absorbance of light that occurs when the NADPH cofactor is consumed by the enzyme (ADH1 or SORD).[45] NADPH, a cofactor, must be present in the solution for the reaction to occur. The raw data is converted into Michaelis-Menten graphs to determine k_{cat} and K_m. For clarity, k_{cat} represents the processing speed of the interaction, the number of substrate molecules that each enzyme site converts per second. Whereas, the K_m approximates the strength of the binding, which is the free substrate concentration when the reaction velocity is half its maximum value or ½ Vmax. Another term of relevance is Vmax, reaction rate when enzyme is fully saturated by substrate.

[45] Danson M, Eisenthal R (2002). *Enzyme assays: a practical approach*. Oxford [Oxfordshire]: Oxford University Press. ISBN 978-0-19-963820-8.

Step by Step Methodology:

Cloning:

Murine *Adh1* and murine *Sord* cDNA were purchased from GE Dharmacon. They were separately cloned into pet 28(+) vectors, using 50 μg/ml kanamycin, with His-6 tag and EcoR1 & XhoI restriction sites, as previously described in Schlicker, et al., 2019. GoTaq (Promega) kit was used for PCR. The plasmids were expressed in BL21 (DE3) cells. A colony was grown in LB broth, with 50 ug/ml kanamycin and incubated overnight at 37 °C, @255 RPM. Colonies were picked from these plates and DNA isolated from these colonies was analyzed using Sanger sequencing. These colonies were picked individually and grown in a 4mL tube overnight at 37°C, @255 RPM. One 4ml bacterial culture was used to inoculate 1L LB overnight at 16°C, @255 RPM with 50 μg/ml kanamycin. The culture was then induced with 750 μM IPTG and grown to an optical density (OD) of 0.6-0.8. The bacterial culture is spun down (7,000 RMP @4 °C) overnight and cell pellets were stored at -80 °C.

Purification:[46]

To prepare the frozen cell pellet for purification, re-suspension and lysing occurred. The cell pellet was prepped on ice with 10 mM Tris, pH 8.0, 100 mM NaCl, and 100 μM PMSF and lysed four times via a French press, as previously described in Anguera, et al, and any changes are listed herein.[47] The pellet supernatant was then run through a vertical purification column containing a maximum of 5 mL of Co^{2+} resin equilibrated with wash buffer (10 mM Tris, pH 7.0,

[46] Anguera, Montserrat C., et al. "Cloning, Expression, and Purification of 5,10-Methenyltetrahydrofolate Synthetase from Mus Musculus." *Protein Expression and Purification*, vol. 35, no. 2, June 2004, pp. 276–283, https://doi.org/10.1016/j.pep.2004.02.010. Accessed 19 Oct. 2020.
[47] Anguera, Montserrat C., et al. "Cloning, Expression, and Purification of 5,10-Methenyltetrahydrofolate Synthetase from Mus Musculus." *Protein Expression and Purification*, vol. 35, no. 2, June 2004, pp. 276–283, https://doi.org/10.1016/j.pep.2004.02.010. Accessed 19 Oct. 2020.

100 mM NaCl, and 12.5 mM imidazole) once. The column was washed with 5 column volumes (approx. 75 mL) of wash buffer before being eluted. The protein was eluted using 5 mL wash buffer. The eluted product was incubated for 2 hours with biotinylated thrombin to cleave the "6xHis" tag, and then thrombin was removed with the Novagen Thrombin Cleavage Capture Kit. The protein fractions and final eluted product was run on an SDS-PAGE gel, stained with Coomassie blue to ensure successful purification. Dialysis was conducted overnight at 4°C in a solution of 50 mM Hepes, pH 6.8, 10 mM β-mercaptoethanol, and 30% glycerol. A Lowry assay was used to assess enzyme concentration for kinetic analysis.[48]

Kinetic activity of enzyme was assessed with a spectrophotometer, by monitoring A340 upon oxidation of NADH at room temp. A cuvette contained the assay of purified recombinant enzyme at concentrations (0μg/μl, 2μg /μl, 4μg /μl, 6μg /ul, 8 μg/μl, 10μg/μl), varying erythrose concentrations (for SORD, [erythrose] = 0mM, 200mM, 400mM, 600 mM; for ADH1, [erythrose] = 0mM, 100mM, 200mM, 300 mM) and NADPH buffer (50 mm sodium phosphate buffer, pH 6.8, 200 μm NADPH (or NADH)) at room temperature.

For SORD and ADH1, the *Km*, and V_{max} values from each independent experiment was determined by fitting the data to $Y = V_{max} \times X/(K_m + X)$ where Y represents reaction velocity and X indicates substrate concentration using GraphPad Prism software, as shown in Schlicker et al., 2019. The k_{cat} values were determined by multiplying V_{max} by the enzyme concentration. ADH1 and SORD spectrophotometer measurements were conducted in triplicate. Data is reported as the average of the three trials.

[48] Bensadoun, André, and David Weinstein. "Assay of Proteins in the Presence of Interfering Materials." *Analytical Biochemistry*, vol. 70, no. 1, Jan. 1976, pp. 241–250, https://doi.org/10.1016/s0003-2697(76)80064-4. Accessed 27 Apr. 2020.

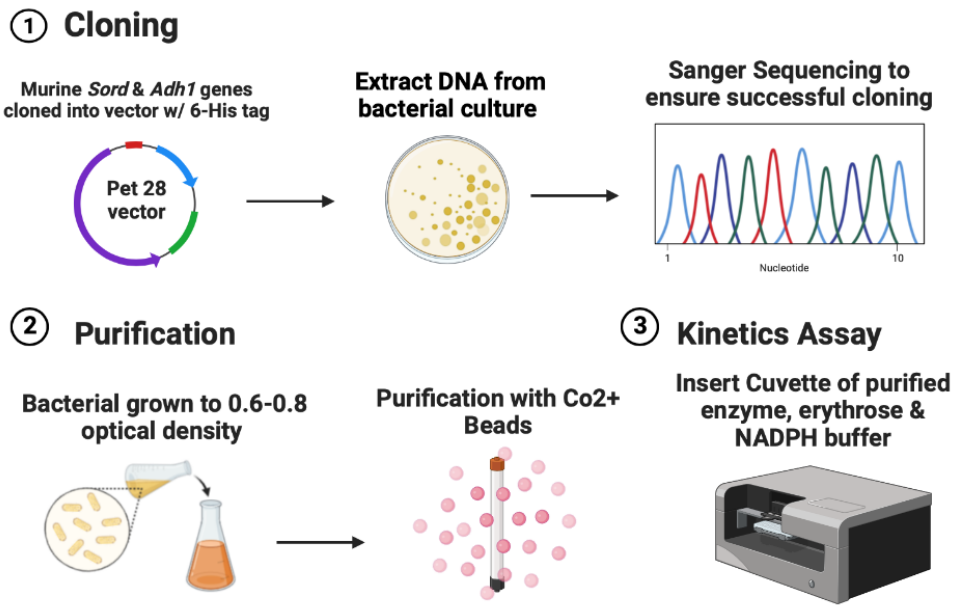


Figure 3: A overview of the methodology of Cloning, Purification and Kinetics Assay used. Made in BioRender.

Part 2: Searching for Isoenzymes of ADH1 & SORD

Background:

As aforementioned in the literature review, a homology, specifically a structural similarity[49] search was the next step to determine what alternate enzymes might also catalyze the final step in the conversion of glucose to erythritol.

This structural similarity search for homologs was conducted computationally with the Research Collaboratory for Structural Bioinformatics (RCSB) Human Protein Data Bank, as it has a digital collection of identified protein structures and variants openly accessible to researchers. It also allows for a refined search based on experimental techniques, (ie. x-ray crystallography, electron microscopy, NMR spectroscopy) used to determine protein structure, the enzyme classification, and species that the protein was isolated from. It also allows the search to be relaxed or constricted with structural motifs of interest in the database. Human ADH1 and SORD have corresponding PDB or protein data bank codes, that are used to identify the protein of interest. The PDB codes are ADH1 (1HSZ) and SORD (1PL7). These codes were searched for with the refinements of the oxidoreductase family, x-ray crystallography as the experimental method, and limited to human and murine species.

Pymol, is a molecular graphic tool that allows for the 3-D visualization, manipulation, and comparison of molecules. You can use python, a popular coding software, to overlay and rotate molecules with various coding commands. The molecules were overlayed, rotated and observationally analyzed to determine homology. A mathematical analysis (RMSD) was also generated from pymol to determine the degree of overlap. RMSD is the root mean square

[49] Koonin, Eugene V. (2005). Orthologs, Paralogs, and Evolutionary Genomics 1. https://doi.org/10.1146/annurev.genet.39.073003.114725

deviation, which is "is a measure of the degree of structural overlap between two protein structures and is an average measure of how far apart an α-Carbon (Cα) atom is in one structure is from the Cα atom from the other structure after they are aligned."[50]

Step-by-Step Methodology

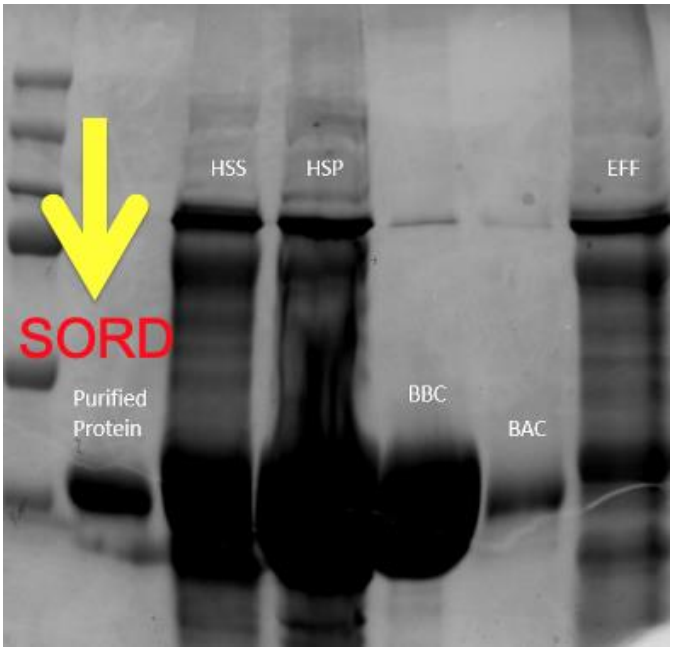
A model of ADH1 with bound NAD^+, PO^{4-} and Zn^{2+} was constructed from PDB entry 1HSZ. A model of SORD with NAD+, n-arylpiperazine, and Zn^{2+} was constructed from PDB entry 1PL6. Structural homologs were identified computationally via a search through previously identified protein structures found in the RCSB Human Protein Data Bank. The researcher used the advanced search builder tool and set the settings to relaxed. This allows for a wider range of potential structural motifs. The refinements were as follows: Scientific name of Source of Organisms (Homo Sapiens & Mus Musculus), Experimental method (X-ray Diffraction) and Enzyme Classification Name (Oxidoreductase). Each query search was compared to the constructed model with pymol. Each structure was overlayed to determine structural similarity. The command 'fetch' was used to pull the structures from the PDB database. The command 'remove resn hoh' was used to remove waters to clarify the structures. The command 'align [PDB CODE #1] chain a, [PDB CODE #2] chain a]' was used to compare the PDB codes ADH1 (1HSZ) & SORD (1PL7), respectively, to other PDB codes of relevance. The structures were aligned, re-aligned, and rotated to ensure that there was observable structural overlap or no overlap. RMSD values were considered after an observational overlap was seen, the overlap ranges accepted were between 2.0 Å and 2.5 Å.[51]

[50] Amgen Scholars. "Aligning Three-Dimensional Structures in PyMOL." *LabXchange*, https://www.labxchange.org/library/items/lb:LabXchange:6d307d41:html:1.
[51] Kwon, Sohee, and Chaok Seok. "CSALIGN and CSALIGN-Dock: Structure Alignment of Ligands Considering Full Flexibility and Application to Protein–Ligand Docking." *Computational and Structural Biotechnology Journal*, vol. 21, 2023, pp. 1–10., https://doi.org/10.1016/j.csbj.2022.11.047.

Results:

Part 1: Determining Enzymatic Activity

ADH1 and SORD were successfully cloned and purified (Figure 4a,b). The gel was stained with Coomassie brilliant blue, which stains the protein fractions throughout the purification process. The respective bands and corresponding enzymes are labeled. As indicated in Figure 4a,b, there is only one protein band in either of the purified protein lanes for both SORD and ADH1, indicating that the protein isolated was pure and not contaminated with other bacterial proteins.



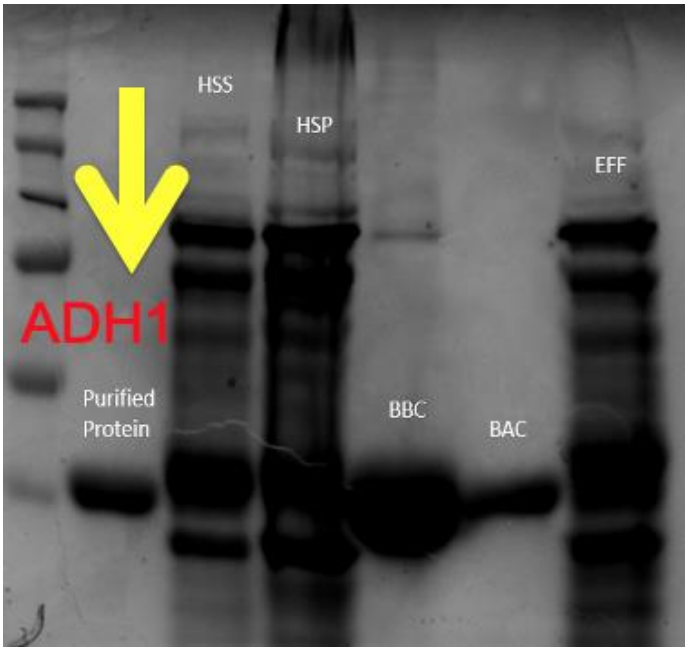


Figure 4a,b: Purification gels with purified recombinant enzymes of interest.

4a: SORD (sorbitol dehydrogenase)

4b: ADH1 (alcohol dehydrogenase)

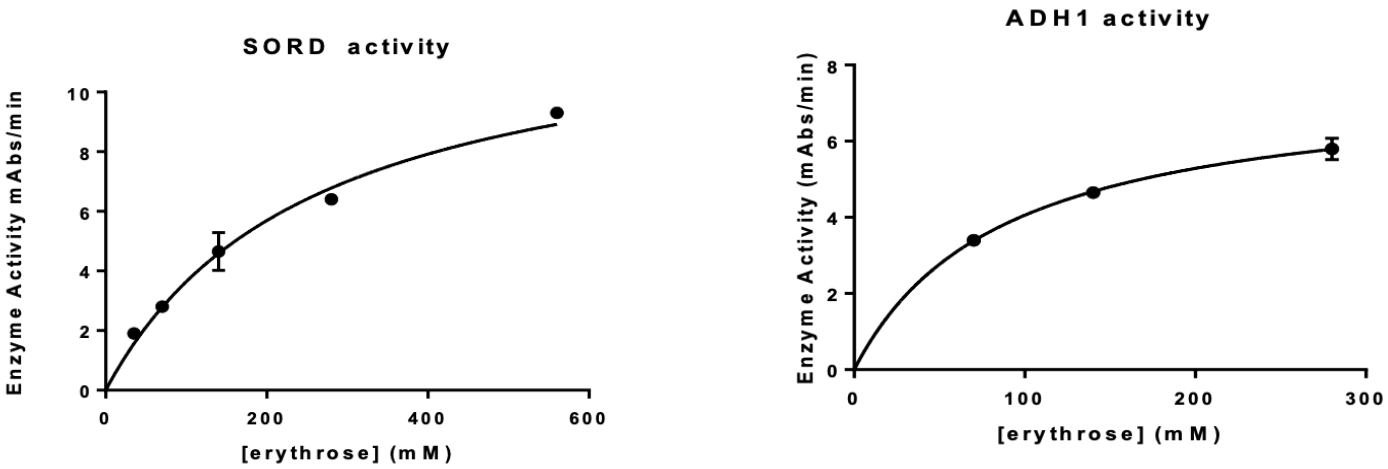


Figure 5a,b: Sample Michaelis–Menten Kinetics Graph from Recombinant Mouse SORD (sorbitol dehydrogenase) and ADH1 (alcohol dehydrogenase)

To determine enzymatic activity, a Michaelis-Menten kinetics graph of recombinant murine ADH1 and SORD enzymes was generated. The graph (Figure 5a,b) shows that there is recombinant murine enzymatic activity present *in vitro* for both ADH1 and SORD in catalyzing the final step in the conversion of glucose to erythritol. The kinetics assays were repeated several times, and summarized data is shown in Table 1.

Enzyme & Species	Km (mm)	Vmax (mABS/min)	Kcat (min-1)
Mice ADH1	137.0 ± 34.8	14.3 ± 2.0	2.6 ± 1.7
Human ADH1	486.0 ± 19	17 .0 ± 0.5	71 ± 5.0
Mice SORD	325.5 ± 60.2	71.0 ± 14.5	28.3 ± 12.2
Human SORD	118.0 ± 37	10 ± 2	95 ± 21.0

Table 1: Sample Michaelis–Menten Kinetics from Recombinant Mouse SORD & ADH1 compared to Human Recombinant SORD (sorbitol dehydrogenase) and ADH1 (alcohol dehydrogenase). Results from human recombinant enzymes are taken from previously published work (Schlicker, 2019).

The K_m , V_{max}, k_{cat} values were fit to the data using Michaelis-Menten equations with the assistance of GraphPad Prism software. There is enzymatic activity of both murine ADH1 and SORD, and the activity is lower than the respective human homologs. Lower activities could be caused by a variety of factors, but because the activity is present in the recombinant enzymes, future murine experiments are likely to inform on the role of these enzymes in metabolism of glucose to erythritol. In summary, these data suggest that ADH1 and SORD knockout mouse models are suitable models for future studies designed to uncover the characteristics of the mammalian erythritol biosynthesis pathway.

Part 2: Computational Homology Study

After mouse models lacking ADH1 or SORD enzymes revealed the unexpected result of little to no change in erythritol production, a computational search for alternate enzymes catalyzing the activity was conducted. This computational approach is hypothesis-generating, not to replace wet lab experimentation. It is important to note that the computational search encompassed both human (Homo Sapiens) and mouse (Mus Musculus), as we know both enzymatic activities in murine and human samples. All of the alternate enzymes investigated were in the oxidoreductase family and the enzyme structures used were determined with x-ray crystallography.

There were 4 results for the ADH1(PDB:1HSZ) homology search. All found results were investigated for structural similarity using a computational software called pymol. Three different ADH1 homologs were identified as complimentary and suitable for future investigation given their structural similarity to ADH1. The PDB numbers of greatest interest, 1M6H and 1M6W were both ADH5 variants. ADH5 is an enzyme that is in the class III alcohol dehydrogenase family.[52] The only difference between the two alternate enzymes is their ligand groups. PDB: 1M6W has an additional Zn^{2+} group on top of the two ligand groups it has in common with PDB: 1M6H (2 two K^{+} ions and three PO^{4-} groups). A sample figure, Figure 6, shows considerable structural overlap between the original ADH1 structure and its ADH5 homolog.[53]Another homolog, PDB: 1E3E was identified and had a strong overlap, however, it

[52] "ADH5 Orthologs." *NCBI*, www.ncbi.nlm.nih.gov/gene/128/ortholog/?scope=89593. Accessed 6 Mar. 2023.
[53] University of Waterloo. "Protein Structural Overlap." *Waterloo*, student.cs.uwaterloo.ca/~cs483/CourseNotes/05_Protein_Structure_Overlap_H.pdf.

was not of interest because it is an ADH2/ADH4 variant. ADH4 was not significantly shown to reduce erythritol production in previous cell culture experiments.[54]

Table 2 gives a summary of the significant homology findings and provides the ligands that differ from the original ADH1 structure used in the search. These ligands may impact enzymatic activity and will be relevant for future research, hence they are included in the table summary.

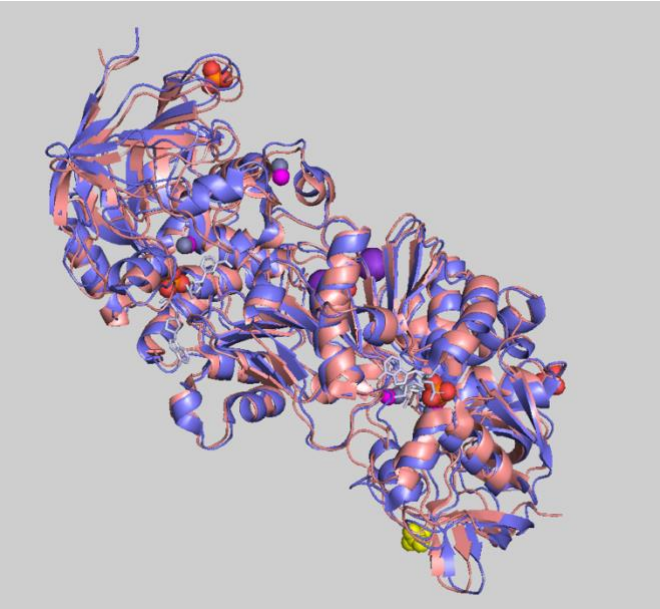


Figure 6: Sample figure for homology search, ADH1 PDB: 1M6H (purple) overlap with ADH5, PDB: 1HSZ (pink).

[54] Semira Ortiz, Doletha Szebenyi, Martha Field. Endogenous Synthesis of Erythritol, a Novel Biomarker of Weight Gain (P15-016-19). Current Developments in Nutrition,Volume 3, Supplement 1, 2019.

PDB Code:	Gene Name:	Isozyme Name:	Attributes:
1HSZ (homo sapiens)	ADH1 *	Human Beta-1 Alcohol Dehydrogenase	Zn^{2+} , $P0^{4-}$, and NAD^{+}
1M6H (homo sapiens)	ADH5	Human glutathione-dependent formaldehyde dehydrogenase	2 K^{+} ions and 3 $P0^{4-}$ groups
1M6W (homo sapiens)	ADH5	Human glutathione-dependent formaldehyde dehydrogenase	Zn^{2+} ,2 K^{+} ions and 3 $P0^{4-}$ groups
1E3E (mus musculus)	ADH2, ADH4	Mouse class II alcohol dehydrogenase complex with NADH	Zn^{2+} , NAI

Table 2: ADH1 alternate enzymes investigation. ADH5, PDB: IM6W & 1M6H both represent a form of ADH5 that was used for additional structural analysis. ADH1, PDB: 1HSZ was the original structure for homology analysis. ADH2/ADH4 is another identified alternate enzyme that was not considered due to previous experiments ruling this alternate enzyme out. ADH1 denotes the base structure that is compared to the other homologs.*

There were 169 results for SORD, PDB: 1PL6 homolog search. All 169 possible variants were analyzed with Pymol as aforementioned above. None of the listed isoforms structurally overlapped the SORD (PDB:1PL6), which suggests that there was not significant structural homology. Unlike ADH1, there was no alignment of the major chains. Figure 7 shows SORD structural disagreement or lack of structure overlap between two SORD homologs. It is compared to Figure 6, which shows ADH1 agreement, or clear structural overlap between ADH1 homologs.

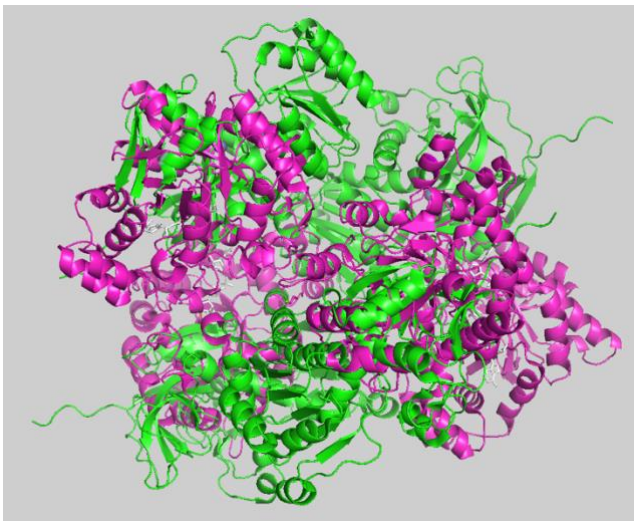


Figure 7:

An overlap between SORD, PDB:1PL7 (green) and Alpha Hydroxysteroid PDB:2HE5 (magenta).

Discussion:

A March 2023 paper in Nature Medicine, once again linked elevated serum erythritol to cardiovascular disease development and generated significant interest regarding the use of erythritol as a biomarker (and non-nutritive sweetener) in the popular press.[55] The work presented here contributes to a body of literature that is focused on understanding how endogenous erythritol production is related to the development of cardiometabolic disease. Understanding endogenous erythritol synthesis is an exciting area of study because of its powerful predictive value for cardiometabolic disease onset. Ultimately elevated serum erythritol could be a way to determine disease risk well before an individual suffers from disease onset and risks permanent metabolic or cardiovascular damage.

Most metabolic biomarkers are not able to determine risk prior to disease onset and the development of potentially life-threatening comorbidities. Popular biomarkers such as elevated triglycerides and low-density lipoproteins (LDL) are metabolic lipid biomarkers, which can signal that cardiovascular disease has or is currently occurring.[56] These biomarkers are useful for disease diagnosis, but the individual may already exhibit irreversible damage to the heart and vascular system –lipid biomarkers offer no predictive power. This is what makes erythritol such a promising avenue for biomarker research, as it has the ability to predict cardiometabolic dysfunction up to 20 years prior to onset.[57] As aforementioned in the literature review section, the public health benefit of this measure is essential, given that cardiometabolic diseases make

[55] Witkowski, M., Nemet, I., Alamri, H. *et al.* The artificial sweetener erythritol and cardiovascular event risk. *Nat Med* (2023). https://doi.org/10.1038/s41591-023-02223-9
[56] McGranaghan P, Kirwan JA, Garcia-Rivera MA, Pieske B, Edelmann F, Blaschke F, Appunni S, Saxena A, Rubens M, Veledar E, Trippel TD. Lipid Metabolite Biomarkers in Cardiovascular Disease: Discovery and Biomechanism Translation from Human Studies. Metabolites. 2021 Sep 14;11(9):621. doi: 10.3390/metabo11090621. PMID: 34564437; PMCID: PMC8470800.
[57] Rebholz, et al., 2018

up the vast majority of deaths globally.[58] Overall, erythritol is one of the many metabolic biomarkers that have been discovered, and given the unique properties outlined above, it is vital that we continue to understand the biosynthesis of endogenous erythritol production so that it can be utilized in clinics as a predictive biomarker.

On top of this powerful predictive power, biomarkers are only as useful as they are cost-effective, accessible and rapid.[59] Erythritol can be measured in blood plasma, making it a relatively non-invasive and accessible biomarker for patients and practitioners. A blood plasma erythritol test could be folded into traditional blood workups, such as a Basic Metabolic Panel (BMP) that analyzes blood sugar levels or a Comprehensive Metabolic Panel (CMP) that examines blood sugar levels and liver function. With Medicaid, BMP and CMP exams costs only around $15.[60] For reference, other blood-based biomarkers, like troponin, cost around $12 on Medicaid[61], and they can only inform practitioners that a patient is suffering from acute heart damage, rather than predict that this damage may occur in the future.[62]As for biomarkers that cannot be easily taken from blood, they are far more invasive, expensive and inaccessible. Collecting a biopsy or spinal tap to gather tissue or CSF (cerebral spinal fluid), respectively, is far more invasive and expensive.[63]

---

[58] World Health Organization. "The Top 10 Causes of Death." *World Health Organization*, 9 Dec. 2020, www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
[59] Gersh, B.J. "Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death." *Yearbook of Cardiology*, vol. 2008, Jan. 2008, pp. 306–308, https://doi.org/10.1016/s0145-4145(08)04037-9.
[60] MD, Frank DiVincenzo. "How Much Does Blood Work Cost in 2022?" *K Health*, 9 Mar. 2022, khealth.com/learn/healthcare/how-much-does-bloodwork-cost/.
[61] "CodeMap®-Medicare Reimbursement Information." *Www.codemap.com*, www.codemap.com/abbott/userhome.cfm?category=24&desc=Cardiac%20Markers. Accessed 6 Mar. 2023.
[62] Mount Sinai. "Troponin Test Information | Mount Sinai - New York." *Mount Sinai Health System*, www.mountsinai.org/health-library/tests/troponin-test.
[63] Mayeux, Richard. "Biomarkers: Potential Uses and Limitations." *NeuroRX*, vol. 1, no. 2, Apr. 2004, pp. 182–188, www.ncbi.nlm.nih.gov/pmc/articles/PMC534923/, https://doi.org/10.1602/neurorx.1.2.182.

It is also important to consider how rapidly biomarkers can be measured. Biomarkers that can be assayed in a few hours in a local lab will be far more accessible than those that require advanced equipment and specialized practitioners. This is one avenue where erythritol as a biomarker falls short. Currently, the actual erythritol blood plasma analysis requires a GC-MS[64] or a ultra-performance liquid chromatography[65] machine for measurement. This equipment is specialized and can be found in many university research labs, but less so in hospitals (especially those that are in rural and medically underserved communities).[66] However, with additional research and attention being brought to erythritol, and biomarkers generally[67], it is possible that an efficient, lower-tech assay will be created.

Specifically, the purpose of this work was to better understand the reductive enzymes in the erythritol biochemistry pathway, which is one piece of a larger puzzle that aims to explain the role that erythritol plays in metabolism. To learn more about the endogenous erythritol pathway and its reductive enzymes, a whole animal system was needed to determine how ADH1 and SORD activities contributed to erythritol production. However, mouse models lacking these enzymes did not display reduced erythritol synthesis,[68] indicating that either these enzymes are not rate-limiting of erythritol synthesis or that there are additional enzymes that catalyze this activity. As shown in Table 1, lower enzymatic activity was observed in the murine enzymes

[64] Ortiz, Semira R, and Martha S Field. "Chronic Dietary Erythritol Exposure Elevates Plasma Erythritol Concentration in Mice but Does Not Cause Weight Gain or Modify Glucose Homeostasis." *The Journal of Nutrition*, vol. 151, no. 8, 1 Aug. 2021, pp. 2114–2124,.
[65] Rebholz CM, Yu B, Zheng Z, Chang P, Tin A, Köttgen A, Wagenknecht LE, Coresh J, Boerwinkle E, Selvin E. Serum metabolomic profile of incident diabetes. Diabetologia. 2018 May;61(5):1046-1054. doi: 10.1007/s00125-018-4573-7. Epub 2018 Mar 20. PMID: 29556673; PMCID: PMC5878141.
[66] Cyr, M.E., Etchin, A.G., Guthrie, B.J. *et al.* Access to specialty healthcare in urban versus rural US populations: a systematic literature review. *BMC Health Serv Res* **19**, 974 (2019). https://doi.org/10.1186/s12913-019-4815-5
[67] "Biomarkers on a Roll." *Nature Biotechnology*, vol. 28, no. 5, 1 May 2010, pp. 431–431, www.nature.com/articles/nbt0510-431, https://doi.org/10.1038/nbt0510-431.
[68] Elevated plasma and urinary erythritol is a biomarker of excess simple carbohydrate intake in mice Semira R. Ortiz, Martha S. Field bioRxiv 2022.12.04.519026; doi: https://doi.org/10.1101/2022.12.04.519026

compared to the human enzymes. This discrepancy may explain the observed differences between *in vitro* kinetics of the recombinant mouse enzymes and the *in vivo* murine studies, in which erythritol production was not changed as a result of *Adh1* or *Sord* deletion. Another theory is that lower activities might be due to differences in basal metabolic rate between mice and humans; the rate, which accounts for differences in body mass, is seven times greater in mice than in humans.[69] Another explanation for the discrepancy is that there is an alternate enzyme (perhaps in addition to ADH1 and SORD) that catalyzes the final step in the conversion of glucose to erythritol. Thus, a computational homology search was conducted to form new hypotheses about alternate enzymes that may be catalyzing the erythritol biosynthesis pathway.

ADH1 has a variety of variant enzymes that have been studied in depth for ethanol metabolism, but less so for glucose metabolism. Table 2 shows the three alternate enzymes that showed structural overlap. The two most promising candidates based on their structural agreement and novelty were PDB: 1M6H and 1M6W. Both of these enzymes were ADH5. Thus, as explained below in future directions, ADH5 is viewed as a promising candidate for further exploration in cells. The structural overlap can be seen in Figure 6.

As for SORD, there was no found structural agreement of homologs in all 169 results in the RCSB database search. Figure 7a shows the SORD disagreement, this lack of overlap of the original structure PDB: 1PL7 and PDB: 2HE5. Thus, there was no found alternate enzymes of interest that could be used for further investigation. This does not mean that there is not an alternate enzyme of interest, just that a computational means was not able to form a hypothesis

[69] Demetrius L. Of mice and men. When it comes to studying ageing and the means to slow it down, mice are not just small humans. EMBO Rep. 2005 Jul;6 Spec No(Suppl 1):S39-44. doi: 10.1038/sj.embor.7400422. PMID: 15995660; PMCID: PMC1369270.

about one. A computational search is meant to create a hypothesis for future experiments, not to rule them out.

Future Directions:

A future project direction could be to identify and support the potential of ADH5 as an ADH1 isoform that catalyzes the production of erythrose to erythritol. Mouse models lacking ADH1 or SORD did not show a significant decrease in erythritol production[70], despite the purified recombinant murine ADH1 and SORD exhibiting this activity *in vitro* (Table 1). These findings suggests that additional enzymes may catalyze this activity *in vivo*. Therefore, a computational investigation of isoforms with structural similarity and function was conducted. ADH5 was identified as a structural homolog of interest. ADH1 is a class 1 enzyme. ADH5 is a class 3 enzyme. The major difference between class 1 & 3 enzymes, is that class three cannot metabolize ethanol and Class 3 enzymes are a crucial component for the cellular metabolism of formaldehyde.[71] Also of note is that Zn^{2+}, K^{+} and PO^{4-} bind like a ligand to ADH5. ADH1 only has Zn^{2+} and PO^{4-} binding, as well as a co-enzyme NAD^{+}. Interestingly, ADH5 is expressed in human and murine tissues (e.g., kidney & liver),[72] but not in rabbit liver. Rabbit liver was the main experimental model organism where ADH1 and SORD was purified to initially test if recombinant human ADH1 and SORD enzymes were able to catalyze erythrose to erythritol in the seminal Schlicker, 2019 paper.[73] Thus, ADH5 may have been missed in the initial experimental trials that determined ADH1 and SORD to be the two major enzymes catalyzing this pathway.

[70] Elevated plasma and urinary erythritol is a biomarker of excess simple carbohydrate intake in mice
Semira R. Ortiz, Martha S. Field. bioRxiv 2022.12.04.519026; doi: https://doi.org/10.1101/2022.12.04.519026
[71] Pontel, Lucas B., et al. "Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen." *Molecular Cell*, vol. 60, no. 1, Oct. 2015, pp. 177–188, https://doi.org/10.1016/j.molcel.2015.08.020. Accessed 29 Jan. 2021.
[72] "ADH5 Orthologs." *NCBI*, www.ncbi.nlm.nih.gov/gene/128/ortholog/?scope=89593. Accessed 6 Mar. 2023.
[73] Schlicker, et al., 2019

A proposed experiment to test if ADH5 catalyzes the production of erythrose to erythritol will involve a cell culture experiment in which *Adh5* expression is decreased. This experiment could be performed in A549 lung cancer cells as previously described in Schlicker, et al., 2019.[74] The cell line that could be used is A549 with a modification called C2735[75]. This line impairs the function of the Keap1 ubiquitinating protein; without this ubiquitinating, nfr2 expression is increased.[76] Nfr2 is a transcription factor that regulates oxidant expression and regulates oxidation damage. An increase in nfr2 increases oxidative stress, which increases the production of erythritol in cells. This cell line has constitutively active[77] erythritol production, making it suitable for this analysis. A siRNA knockdown silences the expression of an enzyme of interest; in this case[78], *Adh5* expression will be KD. The hypothesis is that this *Adh5* KD will significantly decrease the production of erythritol in A549 cells with C2735 modification compared to the (WT wildtype, no modification) A549 with the C2735 modification.

[74] Schlicker, et al., 2019
[75] Zhang HX, Chen Y, Xu R, He QY. Nrf2 mediates the resistance of human A549 and HepG2 cancer cells to boningmycin, a new antitumor antibiotic, in vitro through the regulation of glutathione levels. Acta Pharmacol Sin. 2018 Oct;39(10):1661-1669. doi: 10.1038/aps.2018.21. Epub 2018 May 10. PMID: 30287928; PMCID: PMC6289326.
[76] Hayes JD, McMahon M. NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. Trends Biochem Sci. 2009 Apr;34(4):176-88. doi: 10.1016/j.tibs.2008.12.008. Epub 2009 Mar 25. PMID: 19321346.
[76] Han H. RNA Interference to Knock Down Gene Expression. Methods Mol Biol. 2018;1706:293-302. doi: 10.1007/978-1-4939-7471-9_16. PMID: 29423805; PMCID: PMC6743327.

Acknowledgments:

I would first like to express my gratitude to Dr. Martha Field, assistant professor in the Division of Nutrition at Cornell University, for her mentorship, technical lab training and support, as well as her dedication to helping me complete the writing of this thesis. Secondly, I would like to thank Semira Ortiz, whose Ph.D. is focused on Erythritol, for her help and guidance throughout my undergraduate career as well as for teaching me important wet lab technical techniques and mouse husbandry. Thirdly, I would like to thank my Cornell University-based thesis committee: Dr. Martha Field, Assistant Professor in the Division of Nutritional Sciences; Dr. Linda Nicholson, Director of Undergraduate Biology and Professor of Molecular Biology & Genetics and Dr. Noah Tamarkin, Associate Professor of Anthropology and Science & Technology Studies for their thoughtful comments, edits, and support. Fourthly, I would like to thank the McNair Scholars program at Cornell University, specifically my advisors Dr. Kristin Dade and Dr. Simon Velasquez for living stipends and conference funding that allowed me to conduct and present this thesis work over several summers and semesters. Lastly, I would like to thank the College of Arts & Sciences at Cornell University for awarding me the Einhorn Grant ($755) and Undergraduate Research Grant ($1,250).

References:

"A1C Test." *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 1 Dec. 2022, https://www.mayoclinic.org/tests-procedures/a1c-test/about/pac-20384643.

"ADH5 Orthologs." *NCBI*, www.ncbi.nlm.nih.gov/gene/128/ortholog/?scope=89593. Accessed 6 Mar. 2023.

Addgene. "Addgene: Protocol - How to Inoculate a Bacterial Culture." *Addgene.org*, 2019, www.addgene.org/protocols/inoculate-bacterial-culture/.

Amgen Scholars. "Aligning Three-Dimensional Structures in PyMOL." LabXchange, https://www.labxchange.org/library/items/lb:LabXchange:6d307d41:html:1.

Anguera, Montserrat C., et al. "Cloning, Expression, and Purification of 5,10-Methenyltetrahydrofolate Synthetase from Mus Musculus." *Protein Expression and Purification*, vol. 35, no. 2, June 2004, pp. 276–283, https://doi.org/10.1016/j.pep.2004.02.010. Accessed 19 Oct. 2020.

Bensadoun, André, and David Weinstein. "Assay of Proteins in the Presence of Interfering Materials." *Analytical Biochemistry*, vol. 70, no. 1, Jan. 1976, pp. 241–250, https://doi.org/10.1016/s0003-2697(76)80064-4. Accessed 27 Apr. 2020.

"Biomarkers on a Roll." *Nature Biotechnology*, vol. 28, no. 5, 1 May 2010, pp. 431–431, www.nature.com/articles/nbt0510-431, https://doi.org/10.1038/nbt0510-431.

"CodeMap®-Medicare Reimbursement Information." *Www.codemap.com*, www.codemap.com/abbott/userhome.cfm?category=24&desc=Cardiac%20Markers. Accessed 6 Mar. 2023.

Crabb DW, Matsumoto M, Chang D, You M. Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. Proc Nutr Soc. 2004 Feb;63(1):49-63. doi: 10.1079/pns2003327. PMID: 15099407.

Curtis L. Fritz, Duc J. Vugia, Chapter 14 - Brucellosis, Editor(s): J. Glenn Morris, Duc J. Vugia, Foodborne Infections and Intoxications (Fifth Edition), Academic Press, 2021,Pages 253-264,ISBN 9780128195192, https://doi.org/10.1016/B978-0-12-819519-2.00016-5.

Cyr, M.E., Etchin, A.G., Guthrie, B.J. *et al.* Access to specialty healthcare in urban versus rural US populations: a systematic literature review. *BMC Health Serv Res* **19**, 974 (2019). https://doi.org/10.1186/s12913-019-4815-5

Danson M, Eisenthal R (2002). *Enzyme assays: a practical approach*. Oxford [Oxfordshire]: Oxford University Press. ISBN 978-0-19-963820-8.

Darain, F., Ban, C., & Shim, Y.-B. Development of a new and simple method for the detection of histidine tagged proteins. Biosensors and Bioelectronics. 2004:20(4);857–863. doi:10.1016/j.bios.2004.03.028.

Demetrius L. Of mice and men. When it comes to studying ageing and the means to slow it down, mice are not just small humans. EMBO Rep. 2005 Jul;6 Spec No(Suppl 1):S39-44. doi: 10.1038/sj.embor.7400422. PMID: 15995660; PMCID: PMC1369270.

Dorota A. Rzechonek, Adam Dobrowolski, Waldemar Rymowicz & Aleksandra M. Mirończuk (2018) Recent advances in biological production of erythritol, Critical Reviews in Biotechnology, 38:4, 620-633, DOI: 10.1080/07388551.2017.1380598.

Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. Alcohol Res Health. 2007;30(1):5-13. PMID: 17718394; PMCID: PMC3860432.

El-Kabbani O, Darmanin C, Chung R P. Sorbitol dehydrogenase: structure, function and ligand design. *Curr Med Chem*, 2004, 11(4):465–476.

Elevated plasma and urinary erythritol is a biomarker of excess simple carbohydrate intake in mice. Semira R. Ortiz, Martha S. Field bioRxiv 2022.12.04.519026; doi: https://doi.org/10.1101/2022.12.04.519026

Erythritol is a pentose-phosphate pathway metabolite and associated with adiposity gain in young adults. Proc Natl.

Garrett, H., Reginald, and Charles Grisham. Biochemistry. Boston: Twayne Publishers, 2008.

Gersh, B.J. "Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death." *Yearbook of Cardiology*, vol. 2008, Jan. 2008, pp. 306–308, https://doi.org/10.1016/s0145-4145(08)04037-9.

Hamed Kianmehr, et al. JAMA Netw Open. 2022;5(4):e227705. doi:10.1001/jamanetworkopen.2022.7705

Hamed Kianmehr, et al. JAMA Netw Open. 2022;5(4):e227705. doi:10.1001/jamanetworkopen.2022.7705

Han H. RNA Interference to Knock Down Gene Expression. Methods Mol Biol. 2018;1706:293-302. doi: 10.1007/978-1-4939-7471-9_16. PMID: 29423805; PMCID: PMC6743327.

Harvard Health Publishing. "Abdominal Obesity and Your Health - Harvard Health." *Harvard Health*, Harvard Health, 20 Jan. 2017, www.health.harvard.edu/staying-healthy/abdominal-obesity-and-your-health.

Hayes JD, McMahon M. NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. Trends Biochem Sci. 2009 Apr;34(4):176-88. doi: 10.1016/j.tibs.2008.12.008. Epub 2009 Mar 25. PMID: 19321346.

Homeostasis." *The Journal of Nutrition*, vol. 151, no. 8, 1 Aug. 2021, pp. 2114–2124.

Hootman KC, Trezzi JP, Kraemer L, Burwell LS, Dong X, Guertin KA, Jaeger C, Stover PJ, Hiller K, Cassano PA.

Karl F. Tiefenbacher, in The Technology of Wafers and Waffles II, 2019

Kokkonen P, Beier A, Mazurenko S, Damborsky J, Bednar D, Prokop Z. Substrate inhibition by the blockage of product release and its control by tunnel engineering. RSC Chem Biol. 2021 Jan 11;2(2):645-655. doi: 10.1039/d0cb00171f. PMID: 34458806; PMCID: PMC8341658

Koonin, Eugene V. (2005). Orthologs, Paralogs, and Evolutionary Genomics 1. https://doi.org/10.1146/annurev.genet.39.073003.114725

Koonin, Eugene V. (2005). Orthologs, Paralogs, and Evolutionary Genomics 1. https://doi.org/10.1146/annurev.genet.39.073003.114725

Kwon, Sohee, and Chaok Seok. "CSALIGN and CSALIGN-Dock: Structure Alignment of Ligands Considering Full Flexibility and Application to Protein–Ligand Docking." *Computational and Structural Biotechnology Journal*, vol. 21, 2023, pp. 1–10., https://doi.org/10.1016/j.csbj.2022.11.047.

Mayeux, R. Biomarkers: Potential uses and limitations. *Neurotherapeutics* **1**, 182–188 (2004). https://doi.org/10.1602/neurorx.1.2.182

McGranaghan P, Kirwan JA, Garcia-Rivera MA, Pieske B, Edelmann F, Blaschke F, Appunni S, Saxena A, Rubens M, Veledar E, Trippel TD. Lipid Metabolite Biomarkers in Cardiovascular Disease: Discovery and Biomechanism Translation from Human Studies.

Metabolites. 2021 Sep 14;11(9):621. doi: 10.3390/metabo11090621. PMID: 34564437; PMCID: PMC8470800.

MD, Frank DiVincenzo. "How Much Does Blood Work Cost in 2022?" *K Health*, 9 Mar. 2022, khealth.com/learn/healthcare/how-much-does-bloodwork-cost/.

Mount Sinai. "Troponin Test Information | Mount Sinai - New York." *Mount Sinai Health System*, www.mountsinai.org/health-library/tests/troponin-test.

Ortiz SR, Field MS. Mammalian metabolism of erythritol: a predictive biomarker of metabolic dysfunction. Curr Opin Clin Nutr Metab Care. 2020

Ortiz SR, Heinz A, Hiller K, Field MS. Erythritol synthesis is elevated in response to oxidative stress and regulated by the non-oxidative pentose phosphate pathway in A549 cells. Front Nutr. 2022 Oct 6;9:953056. doi: 10.3389/fnut.2022.953056. PMID: 36276829; PMCID: PMC9582529.

Ortiz SR, Field MS. Chronic Dietary Erythritol Exposure Elevates Plasma Erythritol Concentration in Mice but Does Not Cause Weight Gain or Modify Glucose Homeostasis. J Nutr. 2021 Aug 7;151(8):2114-2124. doi: 10.1093/jn/nxab130. PMID: 34091676.

Pelley, John W., et al. "4 - Enzymes and Energetics." *Biochemistry*, Mosby/Elsevier, Philadelphia, PA, 2011.

Pontel, Lucas B., et al. "Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen." *Molecular Cell*, vol. 60, no. 1, Oct. 2015, pp. 177–188, https://doi.org/10.1016/j.molcel.2015.08.020. Accessed 29 Jan. 2021.

Rebholz CM, Yu B, Zheng Z, Chang P, Tin A, Köttgen A, Wagenknecht LE, Coresh J, Boerwinkle E, Selvin E. Serum metabolomic profile of incident diabetes. Diabetologia. 2018 May;61(5):1046-1054. doi: 10.1007/s00125-018-4573-7. Epub 2018 Mar 20. PMID: 29556673; PMCID: PMC5878141.

Schlicker L, Szebenyi DME, Ortiz SR, Heinz A, Hiller K, Field MS. Unexpected roles for ADH1 and SORD in catalyzing the final step of erythritol biosynthesis. J Biol Chem. 2019 Nov 1;294(44):16095-16108. doi: 10.1074/jbc.RA119.009049. Epub 2019 Sep 11. PMID: 31511322; PMCID: PMC6827307.

Speakman, J., Hambly, C., Mitchell, S. and Król, E. (2007), Animal models of obesity. Obesity Reviews, 8: 55-61. https://doi.org/10.1111/j.1467-789X.2007.00319.x

Srikanth, KK, and JA Orrick . "Biochemistry, Polyol or Sorbitol Pathways - NCBI Bookshelf." *National Library of Medicine*, 2022.

Sugar Alcohol - Yale New Haven Hospital,https://www.ynhh.org/services/nutrition/sugar-alcohol.

The University of Waterloo. "Protein Structural Overlap." *Waterloo*, student.cs.uwaterloo.ca/~cs483/CourseNotes/05_Protein_Structure_Overlap_H.pdf.

Trigiano, Robert N., and Dennis J. Gray. "Chapter 14 - Techniques for Nucleic Acid Engineering: The Foundation of Gene Manipulation." *Plant Development and Biotechnology*, CRC, Boca Raton, FL, 2004.

Tsao, Connie W., et al. "Heart Disease and Stroke Statistics—2022 Update: A Report from the American Heart Association." *Circulation*, vol. 145, no. 8, 2022, https://doi.org/10.1161/cir.0000000000001052.

Library, UC Berkeley. "Library Guides: NCBI Bioinformatics Resources: An Introduction: NCBI Homepage." *NCBI Homepage - NCBI Bioinformatics Resources: An Introduction - Library Guides at UC Berkeley*, https://guides.lib.berkeley.edu/ncbi.

"Uniprot ." *UniProt*, https://www.uniprot.org/uniprotkb/Q00796/entry.

Wang, Zhe, and Bing Yu. "Response Letter Regarding Article, 'Metabolomic Pattern Predicts Incident Coronary Heart Disease.'" *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 39, no. 8, 2019, https://doi.org/10.1161/atvbaha.119.313013.

Wang, Zhe, and Bing Yu. "Response Letter Regarding Article, 'Metabolomic Pattern Predicts Incident Coronary Heart Disease.'" *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 39, no. 8, 2019,

Watson JD (2007). *Recombinant DNA: genes and genomes: a short course*. San Francisco: W.H. Freeman. ISBN 978-0-7167-2866-5.

Witkowski, M., Nemet, I., Alamri, H. *et al.* The artificial sweetener erythritol and cardiovascular event risk. *Nat Med* (2023). https://doi.org/10.1038/s41591-023-02223-9

World Health Organization. "The Top 10 Causes of Death." *World Health Organization*, 9 Dec. 2020, www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

Zhang HX, Chen Y, Xu R, He QY. Nrf2 mediates the resistance of human A549 and HepG2 cancer cells to boningmycin, a new antitumor antibiotic, in vitro through the regulation of glutathione levels. Acta Pharmacol Sin. 2018 Oct;39(10):1661-1669. doi: 10.1038/aps.2018.21. Epub 2018 May 10. PMID: 30287928; PMCID: PMC6289326.