EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

MATHEMATISCH-
NATURWISSENSCHAFTLICHE
FAKULTÄT

Master's Thesis

Reconstruction of a metabolic model of *Streptococcus sanguinis*

Tobias Fehrenbach

September 28, 2023

Supervisors: Jun.-Prof. Dr. Andreas Dräger and Prof. Dr. Lisa Maier

1 Background and Motivation

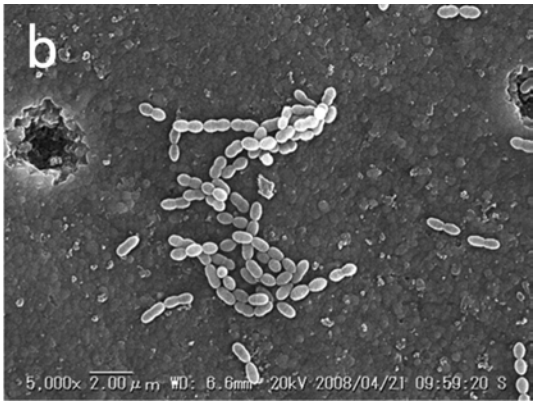

Figure 1 | Adherence of *Streptococcus sanguinis* to saliva-coated hydroxyapatite discs [1].

Commensal bacteria within the human microbiome are usually non-harmful or even beneficial to the individual host. However, perturbations in dietary changes of the host, antibiotic use, genetic transformation of the bacteria itself, or miscolonization within the body can lead to severe dysbiosis, causing commensals to become pathogens [2]. With over 10,000 microbial species within the human microbiome and limited resources for laboratory research, *in silico* approaches to study single species and microbial communities are a cost-effective alternative [3]. One approach is constraint-based stoichiometric modeling. Here, genome-scale metabolic reconstructions are built with a structured knowledge base in a way that the metabolism of an organism is rebuilt mathematically. Such models can then be used to study microbial metabolism in different environments (e.g., oral cavity, gut) and phenotypic reactions to different substances [4, 5, 6]. *Streptococcus sanguinis*, formerly *S. sanguis*, is mainly described as commensal within the humanoid oral cavity, where it is an integral part of the healthy plaque biofilm and antagonizes pathogens by producing hydrogen peroxide [7, 8]. However, the gram-positive, facultative anaerobe member of the Viridians Streptococcus group is also one of the most common causes of infective endocarditis [9]. A particular trait of *S. sanguinis* is a large amount of putative surface proteins. This might be the cause of its role in colonizing the upper respiratory tract and acting as a pathogen [10, 7]. Furthermore, *S. sanguinis* genome contains a large number of carbohydrate transporters that enable the consumption of an extensive range of carbohydrate sources [10]. *S. sanguinis* has the ability for genetic transformation or horizontal gene transfer, enabling fast antibiotic resistance acquisition [11]. In 2007, the genome of *S. Sanguinis* strain SK36 (ASM1420v1) was sequenced for the first time. The circular genome consists of roughly 2.4 Mbp, encoding 2,346 genes of which 2,237 were predicted to be proteins. Among those predicted proteins 248 are hypothetical, meaning they are predicted to exist based on the DNA sequencing data but not experimentally verified or characterized [10, 12]. While the Virtual Metabolic Human (VMH) database lists a metabolic model for strain SK36, this work deals with the creation of a genome-scale metabolic model for strain SK1 (ASM19494v1, ATCC 10556, DSM 20567) [13, 14]. In comparison the genome of SK1 is slightly smaller with 2.3 Mbp, it encodes 2,222 genes with 2,144 predicted proteins and 250 hypothetical proteins [15].

2 Aim and Approach

This thesis focuses on the molecular mechanisms of *Streptococcus sanguinis* growth. The aim is to create a high-quality systems biology model of *Streptococcus sanguinis*. Since there are no available models at the start of the thesis, the steps will follow the standard operating procedure of the Computational Systems Biology (CSB) research group and commonly accepted guidelines [16]. These steps will include (a) analyzing the created model, (b) conducting literature research to identify the bacterium's growth conditions, (c) identifying potential gaps in the metabolic network, (d) defining test cases to check the correctness of the model, (e) running simulations, and (f) aligning the model with available data from experimental investigations.

3 Requirements

(a) Understanding of biochemistry and molecular mechanisms, (b) interest in systems biology, basic knowledge of Python programming, (c) attentively for details, and (d) interest in learning SBML [17], MIRIAM annotations [18], and the SBO [19].

References

[1] K. Hojo et al. "Bacterial Interactions in Dental Biofilm Development". In: *J. Dent. Res.* 88.11 (Oct. 2009), pp. 982–990. ISSN: 0022-0345. DOI: 10.1177/0022034509346811. (Visited on 07/29/2023).

[2] B. Stecher et al. "'Blooming' in the gut: how dysbiosis might contribute to pathogen evolution". In: *Nat. Rev. Microbiol.* 11 (Apr. 2013), pp. 277–284. ISSN: 1740-1534. DOI: 10.1038/nrmicro2989. (Visited on 07/29/2023).

[3] *NIH Human Microbiome Project defines normal bacterial makeup of the body*. Online; accessed 29. Jul. 2023. Aug. 2015. URL: https://www.nih.gov/news-events/news-releases/nih-human-microbiome-project-defines-normal-bacterial-makeup-body (visited on 07/29/2023).

[4] N. E. Lewis et al. "Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods". In: *Nat. Rev. Microbiol.* 10.4 (Feb. 2012), pp. 291–305. ISSN: 1740-1534. DOI: 10.1038/nrmicro2737. (Visited on 07/29/2023).

[5] T. R. Maarleveld et al. "Basic concepts and principles of stoichiometric modeling of metabolic networks". In: *Biotechnol. J.* 8.9 (Sept. 2013), pp. 997–1008. ISSN: 1860-7314. DOI: 10.1002/biot.201200291. (Visited on 07/29/2023).

[6] I. Thiele et al. "A protocol for generating a high-quality genome-scale metabolic reconstruction". In: *Nat. Protoc.* 5.1 (Jan. 2010), pp. 93–121. ISSN: 1750-2799. DOI: 10.1038/nprot.2009.203. (Visited on 07/29/2023).

[7] B. Zhu et al. "*Streptococcus sanguinis* biofilm formation & interaction with oral pathogens". In: *Future Microbiol.* 13.8 (June 2018), p. 915. DOI: 10.2217/fmb-2018-0043. (Visited on 07/24/2023).

[8] L. Zeng et al. "Spontaneous Mutants of *Streptococcus sanguinis* with Defects in the Glucose-Phosphotransferase System Show Enhanced Post-Exponential-Phase Fitness". In: *J. Bacteriol.* 203.22 (Oct. 2021), e0037521. ISSN: 1098-5530. DOI: 10.1128/JB.00375-21. (Visited on 09/26/2023).

[9] S. Paik et al. "Identification of Virulence Determinants for Endocarditis in *Streptococcus sanguinis* by Signature-Tagged Mutagenesis". In: *Infect. Immun.* 73.9 (Sept. 2005), p. 6064. DOI: 10.1128/IAI.73.9.6064-6074.2005. (Visited on 07/24/2023).

[10] P. Xu et al. "Genome of the Opportunistic Pathogen *Streptococcus sanguinis*". In: *J. Bacteriol.* 189.8 (Apr. 2007), p. 3166. DOI: 10.1128/JB.01808-06. (Visited on 07/24/2023).

[11] L. S. Håvarstein et al. "Identification of the streptococcal competence-pheromone receptor". In: *Mol. Microbiol.* 21.4 (Aug. 1996), pp. 863–869. ISSN: 0950-382X. DOI: 10.1046/j.1365-2958.1996.521416.x. (Visited on 07/24/2023).

[12] P. Xu et al. Streptococcus sanguinis *SK36 genome assembly ASM1420v1*. In: NCBI Genome database. accessed 26. Sep. 2023. Feb. 2023. URL: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000014205.1 (visited on 09/26/2023).

[13] S. Magnúsdóttir et al. *GEM of Streptococcus sanguinis strain SK36*. In: Virtual Metbolic Human database. accessed 26. Sep. 2023. May 2019. URL: https://vmh.life/#microbe/Streptococcus_sanguinis_SK36 (visited on 09/26/2023).

[14] S. Magnúsdóttir et al. "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota". In: *Nat. Biotechnol.* 35 (Jan. 2017), pp. 81–89. ISSN: 1546-1696. DOI: 10.1038/nbt.3703. (Visited on 09/28/2023).

[15] D. Muzny et al. Streptococcus sanguinis *SK1 = NCTC 7863 genome assembly ASM19494v1, Direct Submission*. accessed 26. Sep. 2023. NCBI Genome database, Dec. 2022. DOI: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000194945.1. (Visited on 09/26/2023).

[16] M. A. Carey et al. "Community standards to facilitate development and address challenges in metabolic modeling". In: *Molecular Systems Biology* 16.8 (Aug. 2020), e9235. URL: https://www.embopress.org/doi/abs/10.15252/msb.20199235.

[17] S. M. Keating et al. "SBML Level 3: an extensible format for the exchange and reuse of biological models". In: *Molecular Systems Biology* 16.8 (Aug. 2020), e9110. ISSN: 1744-4292. URL: https://www.embopress.org/doi/abs/10.15252/msb.20199110.

[18] N. Le Novère et al. "Minimum information requested in the annotation of biochemical models (MIRIAM)". In: *Nature Biotechnology* 23.12 (Dec. 2005), pp. 1509–1515. URL: http://www.nature.com/nbt/journal/v23/n12/abs/nbt1156.html.

[19] M. Courtot et al. "Controlled vocabularies and semantics in systems biology". In: *Molecular Systems Biology* 7.1 (Sept. 2011), p. 543. URL: https://doi.org/10.1038/msb.2011.77.