UNIVERSIDADE NOVA DE LISBOA

MScBH

MASTER IN
BIOCHEMISTRY FOR HEALTH

Rafaela Almeida Cordeiro

BSc in Biochemistry

Structural and functional studies on enzymes involved in microbial hydrogen sulfide metabolism

September, 2023

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

NOVA MEDICAL SCHOOL

Rafaela Almeida Cordeiro

BSc in Biochemistry

Structural and functional studies on enzymes involved in microbial hydrogen sulfide metabolism

Dissertation presented to obtain the master's degree in
Biochemistry for Health

Supervisor: José A. Brito, *PhD*, ITQB NOVA

Co-supervisor: João Vicente, *PhD*, ITQB NOVA

Jury

President: Dr. Pedro Matias – ITQB NOVA

Opponent: Prof. Dr. Tiago Folgosa Granja, CBIOS – ULHT

Supervisor: Dr. José A. Brito – ITQB NOVA

ITQB-NOVA

September, 2023

Structural and functional studies on enzymes involved in microbial hydrogen sulfide metabolism

ACKNOWLEDGMENTS

Neste caminho que foi o último ano, a jornada tornou-se especial por vossa causa. E ainda que um agradecimento seja pouco, aqui fica.

Ao José, que me orientou neste projeto, quero agradecer de uma forma muito especial, não só pela oportunidade de fazer parte da equipa e de participar no projeto, mas principalmente por me ter dado a conhecer o mundo da investigação de uma forma sincera, inteligente e divertida. Desde o momento em que comecei que foi muito graças a si que me senti em casa, pela paciência, o entusiasmo, os ensinamentos, a força e o profissionalismo. Obrigada pelas horas de risos, pelas piadas, pelas reuniões nos sofás, pela facilidade de comunicação, pela frontalidade e por tudo o que aprendi graças a si. O meu maior e mais sincero obrigada.

Agradeço também ao meu co-orientador, João Vicente, pelos brain storms e break throughs indispensáveis para que este trabalho tenho chegado até aqui.

Agradeço à Professora Margarida Archer por me ter disponibilizado acesso ao seu laboratório, pela oportunidade de fazer parte da sua equipa e pela preocupação e interesse que sempre demonstrou em relação ao meu trabalho, ao longo de todo ano.

Agradeço de forma muito especial à Professora Teresa Catarino, coordenadora do mestrado, que foi durante este ano um apoio essencial. Agradeço a constante presença, o contínuo suporte, as conversas, os desabafos, os risos e a serenidade que transmite sempre (com ou sem *burpies*) e, acima de tudo, a confiança que depositou em mim nestes dois anos.

O seguinte agradecimento é dado à pessoa que tornou os dias bons em dias melhores e os dias maus em memórias prá vida. A ti, Cláudia, que me fizeste crescer. Obrigada pelo apoio gigante, pela forma clara, organizada e tão tua que me ensinaram tanto. Obrigada por teres sido salva-vidas em tantos momentos e por todas as vezes que me deste na cabeça. Acima de tudo obrigada por tornares tudo mais leve, divertido, fantástico e pelas memórias que ficarão pra sempre. E prometo que de hoje em diante, eu uso luvas!

Quero deixar um agradecimento muito especial á Vanessa e ao Diogo por terem sido mentores imprescindíveis neste projeto. Por me guiarem pelas aventuras do laboratório sempre da forma mais tranquila e acolhedora.

Agradeço à Beatriz Moniz, à Mariana Antunes, à Margarida Pedro e à Sara Patão. As colegas que passaram comigo, lado a lado, todos os desafios que são o ser cientista.

Agradeço à Barbara, minha companheira nos meses frios. Porque as nossas duas cabeças valem tanto como um aluno de doutoramento.

O obrigada mais caloroso vai para a minha família. Para a minha mãe, Dra Cristina, pelo amor e a paciência em lidar comigo e por ter posto de lado tudo só para ficar ao meu lado. Para o meu irmão Tomás, futuro Engenheiro Malateco, e para o meu pai, Eng. Vítor, pela preocupação e carinho ao longo deste ano.

Ás minhas batatitas, Rita Carlota, Beatriz Alves e Maria Branco que fizeram do mais longe o mais perto, tornando todos os problemas das quatro teses, quase inexistentes.

Á família que Deus me deu, a minha madrinha, Maria Inês Marques, ao João Sóusa, á Ana Cardoso, ao Diogo Domingos e ao Ricardo Marques pela motivação e força de continuar.

Ao grupo de pessoas especiais que são a Vera Lemos, o Pedro Afonso, a Helena Souza, o David Puga e a Madalena Pernas por estarem sempre preocupados e me puxarem para cima.

E o melhor ficando para o fim, á Figas agradeço por tudo. Pelo amor, os mimos, as conversas, as visitas ilegais, os Happy meals da meia-noite. Sem ti eu não seria.

A todos os que fizeram parte, se não estão na folha, estão no coração.

Obrigada!

"*Não há lugar para a sabedoria onde não há paciência*."

(Santo Agostinho)

ABSTRACT

Hydrogen sulfide has a physiological role in the human organism. The tight control and balance of its metabolism is very important since high concentrations of bacterial-derived H_2S confers them increased resistance to host leukocytes. Infections caused by *Staphylococcus* (*S.*) *aureus* are very common at present and are associated with resistant strains. *S. aureus* has a *cst* operon that encodes proteins responsible for maintaining the tight balance in H_2S metabolism. One of these proteins is SQR, responsible for the first step of H_2S detoxification.

In collaboration with Giedroc's lab, the major aim of this project is to biochemically and structurally characterize the SQR protein from *S. aureus* (*Sa*SQR). To achieve this goal, a protocol "from gene to structure" was developed. Furthermore, the protocols optimized for the wild-type protein will be applied for two mutant proteins (C133S and C344S) aiming at the full mechanistic characterization.

To achieve this goal, new conditions from previously described in the literature were tested to optimize gene expression and different production. Among the different tests, the protocol settled the most favorable condition for expression, using *Escherichia* (*E.*) *coli* ArcticExpress® as host cells, Terrific Broth medium, and induction with 0.25 mM IPTG for 6 hours at 15 °C. Cells were disrupted by sonication. Purification was carried out starting with immobilized-metal affinity chromatography (IMAC), in batch mode, followed by a size exclusion chromatography (SEC).

The protein was submitted to different thermostability assays, namely Thermal Shift Assay (TSA, also known as "Thermofluor"), that allowed us to improve the purification and storage conditions. In the best determined conditions, crystallization trials were performed using different commercially available screens. From all the tested conditions, a crystal was formed in 25% (w/v) PEG 1500 and 0.1 M SPG buffer pH 6.0, at 20 ºC.

Further studies included the application and modulation of the purification conditions for *Sa*SQR mutant variants, enrolling in new and deeper thermostability assays and replication of crystallization conditions and pursue X-ray crystallography diffraction data to determine protein structure.

Key words:

Staphylococcus aureus, antibiotic resistance, hydrogen sulfide (H_2S), *cst* operon, sulfide:quinone oxidoreductase (SQR), X-ray crystallography

RESUMO

O sulfureto de hidrogénio tem um importante papel fisiológico no organismo humano. O equilíbrio do seu metabolismo é muito importante, uma vez que altas concentrações de H_2S de origem bacteriana lhes confere maior resistência aos leucócitos do hospedeiro. As infeções causadas por *S. aureus* são muito comuns, atualmente, e estão associadas a estirpes resistentes. *S. aureus* possui um operão *cst* que codifica proteínas responsáveis por manter o equilíbrio perfeito no metabolismo do H_2S. Uma dessas proteínas é a SQR, responsável pela primeira etapa da desintoxicação do H_2S.

Em colaboração com Giedroc's lab, o principal objetivo deste projeto foi o de caracterizar bioquímica e estruturalmente a proteína SQR de *S. aureus* (*Sa*SQR). Para atingir este objetivo, um protocolo "do gene à estrutura" foi desenvolvido para esta proteína. Além disso, os protocolos otimizados para a proteína WT serão aplicados para duas proteínas mutantes (C133S e C344S) visando a completa caracterização do seu mecanismo enzimático.

Para atingir esse objetivo, novas condições foram testadas para otimizar a expressão de genes e a produção de proteína. De entre os diferentes testes, o protocolo estabelecido para a condição mais favorável inclui expressão, utilizando *E. coli* ArcticExpress como células hospedeiras, meio Terrific Broth e indução com IPTG 0,25 mM durante 6 horas a 15 °C. As células foram lisadas por sonicação. A purificação foi realizada começando por uma cromatografia de afinidade (IMAC), em modo batch, seguida de cromatografia de exclusão molecular (SEC).

A proteína foi submetida a diferentes ensaios de termoestabilidade, nomeadamente Thermal Shift Assay (TSA, também conhecido como "Thermofluor"), que permitiram melhorar as suas condições de purificação e armazenamento. Nas melhores condições determinadas, foram realizados ensaios de cristalização utilizando diferentes condições disponíveis comercialmente. A partir de todas as condições testadas, formou-se um cristal em 25% (m/v) de PEG 1500 e tampão SPG 0,1 M pH 6,0, a 20 ºC.

Em trabalho futuro deve realizar-se a aplicação e otimização das condições de purificação para os mutantes SaSQR, ensaios de termo-estabilidade, a replicação das condições de cristalização e a aquisição de dados de cristalografia de raios X para determinar a estrutura da SaSQR.

Palavras chave: metabolismo de enxofre, resistência antibiótica, expressão e purificação de proteína, cristalização de proteína

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

Amp	Ampicillin
CBS	Cystathionine β-synthase
CSE	Cystathionine γ-lyase
cst	Copper-sensing operon repressor-like sulfur-transferase
CT	Charge transfer (complexes)
CV	Column volume
DNA	Deoxyribonucleic acid
FAD	Flavin adenine dinucleotide
Gent	Gentamicin
IMAC	Immobilized metal affinity chromatography
IUPAC	International Union of Pure and Applied Chemistry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani medium
MOPS	3-morpholinopropane-1-sulfonic acid
MRSA	Methicillin-resistant *Staphylococcus aureus*
MTS	Mercapto-pyruvate sulfur-transferase
MWCO	Molecular weight cut off
NanoDSF	Nano differential scattering fluorescence

OD_{600}	Optical density at 600 nm
OG	Octyl-beta-Glucoside
PCT	Pre-crystallization trial
rpm	Rotations per minute
SPG	Succinic acid, phosphate, glycine buffer
SQR	Sulfide:quinone oxidoredutase
TB	Terrific Broth medium
TSA	Thermal shift assay
T_m	Melting temperature
WT	Wild type

1

INTRODUCTION

1.1 Origin and history of hydrogen sulfide

In the beginning, Earth was bursting with volcanic eruptions and asteroid bombardments. Surrounded by an anoxic atmosphere, Earth was an inhospitable place devoid of life, even in its simplest forms.[1]

These geothermal conditions led to the emission of hydrogen sulfide (H_2S) into the atmosphere, which in turn was dissolved in the existing water and formed sulfidic anions.[2] Hence, H_2S played major roles as an important reactant, catalyst, and sustainable source of energy.[3] H_2S was instrumental in the production of the initial molecular monomers, such as natural amino acids and lipids, which participated in a series of basic reactions that served as the foundation for the emergence of life. [4]

As evolution progressed, phototrophic beings started to release oxygen into the atmosphere.[2] With the accumulation of oxygen and the decrease of sulfur in the atmosphere, sulfur-dependent organisms had to adapt, thereby creating opportunities for the evolution of oxygen-dependent organisms.[5]

In the present time, sulfur-dependent microorganisms can be found in environments analogous to Earth's early formations, such as hydrothermal vents in the deep ocean. Interestingly, traces of H_2S activity also persist in modern higher organisms, although no longer serving as their primary energy source. Instead, H_2S acts as an important regulator and modulator of metabolism and signaling within these organisms. [2,3]

1.2 Chemical properties of hydrogen sulfide

Hydrogen sulfide (H_2S) was discovered in 1777 by the Swedish scientist, Carl Wilhelm Scheele[6], and it was first described as a toxic molecule in 1973[7]. Formerly called hydrosulfuric acid or sulfhydric acid, owing to its acidic properties when dissolved in water, dihydrogen sulfide and sulfane are now the names recommended for H_2S by the IUPAC (International Union of Pure and Applied Chemistry)[8].

With a molecular weight of 34.08 g/mol and a solubility in water of 3.98 g/L (at 20 ºC), hydrogen sulfide is a colorless gas under ambient temperature and pressure. It is a highly flammable, irritant, poisonous and toxic gas, generically associated with the smell of rotten eggs.[9]

Among the six oxidation states of sulfur (Figure 1.1), H_2S is the most reduced form with an oxidation state of −2. It has pK_a values of 7.04 and 11.96 [4,9] and it is highly soluble in water. Thus, at neutral pH, H_2S predominantly exists in solution as the hydrosulfide monoanion (HS^-) and H_2S, with minimal amounts of the sulfide dianion, S^{2-}.

Figure 1.1: **Schematic representation of sulfur's oxidation states**.

Structurally, it bears resemblance to water. However, the bond angle in H_2S is smaller than in water, measuring at ~92º instead of ~104º (Figure 1.2).[10]

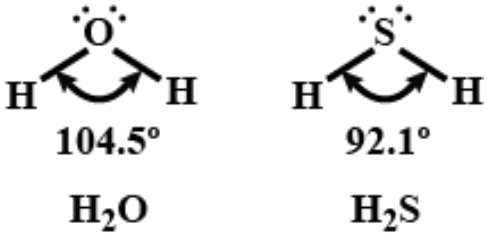

Figure 1.2: **Comparison between the structural geometry of water and hydrogen sulfide**.

Notably, H_2S shares a permeability coefficient akin to other nonpolar dissolved gas molecules like O_2 and CO_2. This characteristic allows it to freely pass through biological membranes and accumulate inside cells, primarily as HS^-.[11,12]

1.3 Homeostasis of hydrogen sulfide

In mammalian cells, H_2S can be synthetized endogenously, or be exogenously captured from the environment. The endogenous synthesis of H_2S requires at least three enzymes: cystathionine β-synthase (CBS), cystathionine γ-lyase (CSE), and mercapto-pyruvate sulfur-transferase (3-MTS). CBS and CSE are cytosolic proteins involved in the trans-sulfuration pathway, responsible for the synthesis of cysteine from homocysteine.[2,13,14] In parallel, 3-MTS converts the 3-mercaptopyruvate derived from cysteine transamination to pyruvate, and transfers sulfur to an acceptor releasing H_2S.[4,13,15]

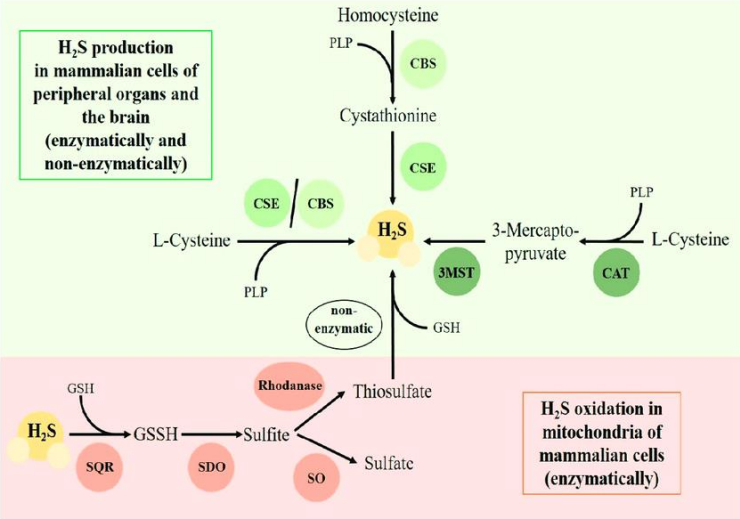

Figure 1.3: **Balance between synthesis and breakdown of H_2S in mammalian cells**. *Adapted from* Denoix *et al.* 2020[16].

The maintenance of H_2S homeostasis, like any other molecule, relies on a strict equilibrium between its synthesis and breakdown (Figure 1.3). For that reason, most cells harbor a highly conserved mitochondrial sulfide oxidation system that produces thiosulfate and sulfate for excretion.[15] Consequently, disturbed H_2S metabolism is associated with numerous human pathologies from cancer to cardiovascular, inflammatory and neurodegenerative disorders.[16]

Since bacteria do not have mitochondria, it could be assumed that they are not able to regulate H_2S stress. However, a mitochondrial-like sulfide oxidation system was found in bacteria and is responsible for the detoxification of H_2S.

This system is regulated by the *cst* (copper-sensing operon repressor-like sulfur-transferase) operon, which is transiently and strongly induced by cellular sulfide stress.[14,15] Transcription of the *cst* operon is regulated by a per- and polysulfide-sensing repressor, the CsoR-like sulfur-transferase repressor (CstR).

The presence and composition of the operon can vary among different bacteria. Moreover, the specific genes and regulatory elements within the operon may differ between species, reflecting variations in the enzymatic steps and regulatory mechanisms involved in cysteine biosynthesis.

1.4 Physiological functions of hydrogen sulfide

Since H_2S initial discovery, its physiological roles have received limited attention. The acceptance of H_2S as an essential regulator in biological processes posed a great challenge. This was due to the lack of knowledge regarding the enzymes responsible for H_2S production, unlike CO or NO which have well-defined enzymatic biosynthesis pathways.[12]

Eventually, H_2S was shown to team up with NO and CO as the third gasotransmitter,[3,12] a small gaseous molecule that transmits information between cells in various parts of the body.[12,15] Gasotransmitters are not stored within secretory vesicles in nerve terminals but, rather, are produced when required.[3]

Hydrogen sulfide is, thus, a biologically important molecule with complex physiological roles, that regulates important functions in the nervous[13], cardiovascular[17], immune, and gastrointestinal systems.[16] (Figure 1.4)

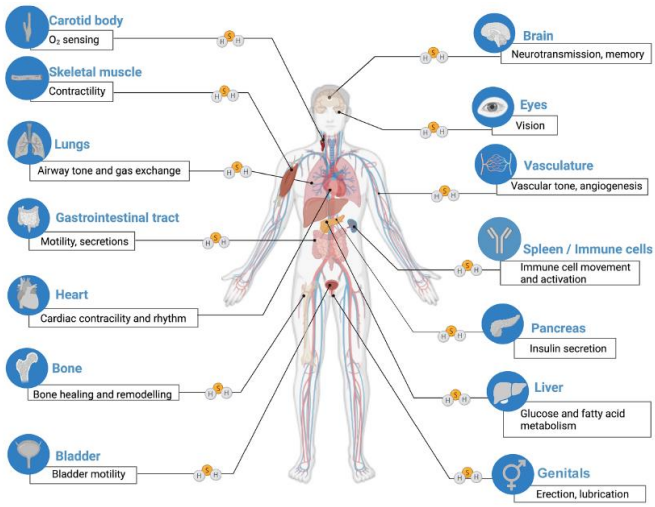

Figure 1.4: **Importance of H_2S across different human physiological systems**. *Adapted from Cirino, 2022.* [16]

The presence of H_2S in the brain was identified in mice[12] and more recently in humans, suggesting a physiological function for this molecule in this organ.[16] In the brain and central nervous system, physiological concentrations of H_2S are beneficial and necessary to regulate some aspects of synaptic activity, as well as being involved in the modification of cognitive function.[13,15]

Recent studies have demonstrated that, in bacteria, high concentrations of H_2S impart increased resistance against leukocyte-mediated immune responses, both *in vitro* and *in vivo*. Although the

underlying mechanisms of this bacterial resistance are unknown, it is theorized that H_2S can prevent oxidative damage by up-regulating the enzymes superoxide dismutase and catalase.

In chemoheterotrophs, organisms unable to synthesize their own organic molecules, such as *S. aureus*, H_2S oxidation is an important source of energy under aerobic conditions, and its cellular levels are tightly regulated through enzymatic activity. The difference in H_2S-tolerance levels across different species is attributed to natural selection according to their natural habitat. This highlights the importance of understanding the intricate interactions between bacteria and their environment, particularly in the context of infectious diseases.

The still growing number of physiological processes regulated by H_2S includes blood flow, cellular stress response, inflammation, immune defense, apoptosis, and energy metabolism.

1.5 Staphylococcus aureus

Staphylococcus aureus is a Gram-positive bacterium that belongs to the micrococcus family and displays a “grape-like” morphology. *S. aureus* has a characteristic golden color, from where it gets the name (*aureus*, Latin for ‘gold’), is a facultative anaerobe and can, therefore, grow aerobically or anaerobically, at temperatures between 18 ºC and 40 ºC.[18] Interestingly, it can be detected by positive results of coagulase, mannitol fermentation and/or deoxyribonuclease tests.[18,19]

Humans are a major reservoir for *S. aureus,* which can colonize different niches within the body including the skin, mucous membranes, nasopharynx, and gut.[18,20] This colonization is facilitated by the bacterium's adaptability to different environments.

1.5.1 Infections and diseases

Noteworthy*, S. aureus* is an opportunistic pathogen that can cause a variety of invasive infections and/or toxin-mediated diseases.[18,21] The severity of *S. aureus* infections depends on the specific strains involved and the site of infection, ranging from minor skin infections to life-threatening diseases, with high morbidity and mortality rates.[19,22]

The spectrum of infections and diseases caused by *S. aureus* includes bacteremia, infective endocarditis, skin and soft tissue infections, osteomyelitis, septic arthritis, prosthetic device infections, pulmonary infections, gastroenteritis, meningitis, toxic shock syndrome, and urinary tract infections.[18]

Colonization with *S. aureus* is a risk factor for healthcare-associated infections in immunocompromised patients and healthcare providers. On the other hand, individuals carrying this pathogen have been observed to exhibit reduced susceptibility to mortality in the context of infections caused by other challenging microorganisms. This suggests that colonization by *S. aureus* may induce beneficial immunological adaptations.[20,23]

S. aureus infections are common both in community-acquired as well as hospital-acquired scenarios and treatment remains challenging due to the emergence of multi-drug resistant strains.[24] Multi-drug resistance in *S. aureus* is very disperse, with the highest resistance levels being observed for penicillin and oxacillin, followed by antibiotics like tetracycline, amikacin, trimethoprim and gentamycin, among others.[25] MRSA (methicillin-resistant *Staphylococcus aureus*) strains are particularly concerning and one of the biggest health concerns in the beginning of the new millennium, being spread around the globe (Figure 1.5).[22,24]

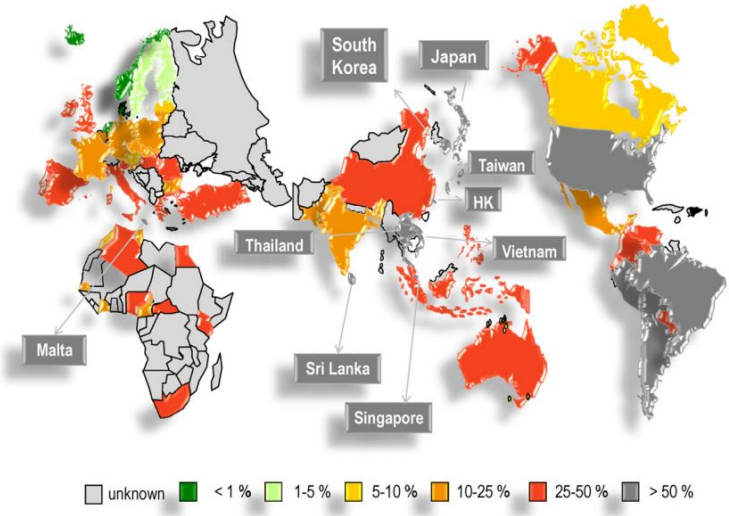

Figure 1.5: **Worldwide prevalence of community-acquired MRSA**. *Adapted from Menezes Alencar (2015).*[24] Portugal has a prevalence of community adquired MRSA between 25-50%.

Among the reasons for this are its increasing incidence, and the existence of two epidemiologically, phenotypically, and genotypically different strains: the community acquired and the health care acquired strains, associated with severe infections with high mortality and morbidity rates.[26]

1.5.2 The *cst* operon in *Staphylococcus aureus*

Recent data reveals that the *cst* operon in *Staphylococcus* encodes a nearly complete mitochondrial-like sulfide oxidation system.[27] The *cst* operon has been observed to be duplicated in the genomes of some virulent and major antibiotic-resistant *S. aureus* strains, such as MRSA (Figure 1.6). This genomic data is consistent with the hypothesis that the sulfur assimilation regulon confers advantage to *S. aureus* during host infection. Hence, the *cst* operon is hypothesized to be involved in bacterial defense during host infection, and has been implicated in the formation of biofilms, an important virulence factor, although this is not well understood yet.[28,29] Indeed, H_2S has been found to act as a protective shield against antibiotic-induced damage, and several studies have shown that H_2S produced by bacteria can modulate immune responses. [27,28]

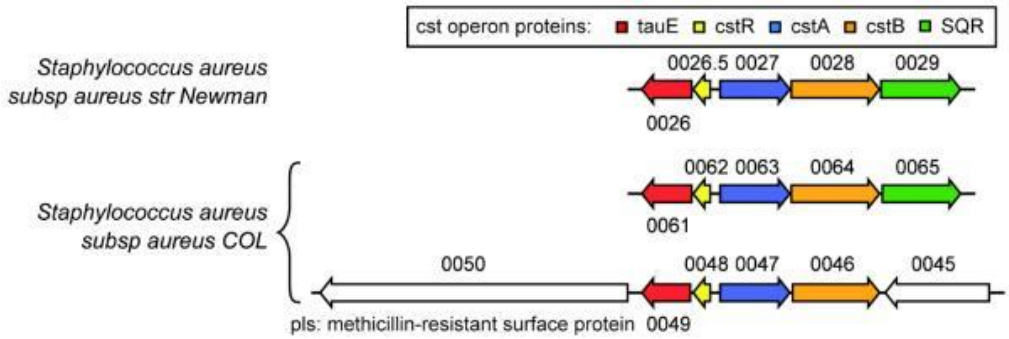

Figure 1.6: **Comparison of the *cst* operon structures** between a nonpathogenic *S. aureus* strain (Newman) and an antibiotic-resistant *S. aureus* strain (COL). *Adapted from Shen, et al (2015)* [30]

In *S. aureus*, the *cst* operon encodes a set of core genes: *tauE*, *cstR*, *cstA* and *cstB*, and *sqr*.[30] Although the physiological function of the *cst* operon is not known[29], the *cstA*, cstB and *sqr* genes have been shown to be necessary to mitigate the effects of cellular sulfide toxicity[28]: its oxidizes sulfide (S^{2-}) to thiosulfate ($S_2O_3^{2-}$) which can be assimilated by this organism.[27–29]

Two possible mechanisms have been hypothesized for the activation of the *cst* operon in *S. aureus*. One, the *cst* genes are induced by any condition that specifically alters thiol-disulfide homeostasis, including, but not limited to, sulfite stress. Another possible inducer of *cst* expression is taurine, a sulfonated non-proteinogenic amino sulfonic acid. Since *S. aureus* does not encode a taurine repressor, this opens the possibility for it to be sensed as a sulfide metabolite.

From the core genes expressed by this operon, the most important are *cstA*, *cstB* and *sqr* (Figure 1.7*)*. From the proteins encoded by these genes, CstA and CstB are well characterized while SQR is not.

CstA has recently been characterized as a multidomain sulfur-transferase that reacts with a persulfide formed on the cysteine desulfurase SufS, as well as with inorganic polysulfides such as sodium tetrasulfide.[30,31] It plays a fundamental role in protecting *S. aureus* against hydrogen sulfide toxicity.[31]

SQR and CstB, on the other hand, are predicted to comprise a system analogous to the first two steps in human mitochondrial H_2S oxidation, starting with the oxidation of H_2S to glutathione persulfide and, subsequently, to sulfite in the presence of molecular oxygen while regenerating reduced glutathione.[14,30]

The CstB protein is predicted to be a three-domain enzyme containing an N-terminal metallo-β-lactamase-like (MBL) domain, a nonheme Fe(II)-containing PDO domain, followed by a pseudo-rhodanese homology (RHD) domain, and a conventional C-terminal rhodanese (Rhod) domain.[30]

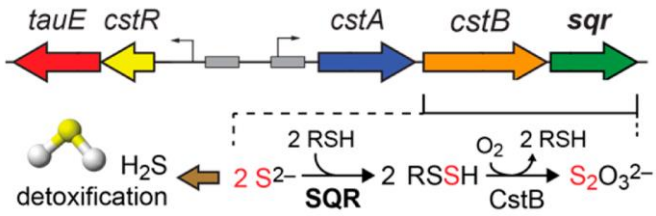

Figure 1.7: **Structure of the *cst* operon in *S. aureus* and function of the encoded proteins in H_2S detoxification**. *Adapted from Shen et al. (2016)* [27]

In short, both CstA and CstB have been biochemically characterized as multidomain proteins, the former being a sulfurtransferase and the latter a non-heme Fe(II) persulfide dioxygenase, and these findings support an H_2S detoxification model in *S. aureus*.[27,30,31]

The initial step of H_2S oxidation in *S. aureus* involves the sulfide:quinone oxidoreductase enzyme SQR, that catalyzes a two-electron oxidation of H_2S to sulfane sulfur (S^0) using the membrane quinone as an electron acceptor.

1.6 Sulfide:quinone oxidoreductases

Sulfide:quinone oxidoreductases (SQR, EC1.8.5.4)[32] are one of the best characterized enzymes to date and play a crucial role in the oxidation of sulfide. These enzymes belong to the flavin disulfide reductase (FDR) family, and their molecular weight ranges from 48 to 55 kDa. SQRs function primarily as dimers although they can also be active as monomers or trimers. [33]

The location of SQR proteins in the cell remains unclear and may differ depending on the organism. In some cases, SQR proteins are found tightly bound to the membrane, which makes their isolation difficult. In other cases, these proteins are associated with the periplasmatic side of the membrane, being easily detached by the addition of high salt concentrations.[3,34] There are also cases where the sulfide oxidation site of SQR is suggested to be located in the cytoplasmatic side of the thylakoid membrane. In any of these cases, it is certain that SQRs are membrane-bound flavoproteins.[3]

These enzymes are essential for the detoxification of H_2S, contributing to cellular sulfide homeostasis. As oxidoreductases, SQRs react with sulfur-containing molecules, in this case sulfide, and transfer two electrons from the sulfide group to a quinone group (the membrane acceptor).

SQRs are found in organisms that thrive in both sulfide-rich and sulfide-deprecated environments suggesting these enzymes perform sulfide oxidation depending on the organism's demands. SQR genes are present in all domains of life having been found in archaea, bacteria, fungi, insects, and vertebrates. Interestingly, SQRs have not yet been identified in plants. SQRs are present in 22% of the 2177 species genomes available to this date (as a single copy or more than one per organism), as

analyzed by Sousa *et al*.[34] and some of these SQRs have been isolated and biochemically characterized.

1.6.1 Historical perspectives on SQRs

The history of SQRs starts in 1975, when Knaff and Buchanan[35] first identified the presence of quinones in the sulfide oxidation of the green sulfur bacterium *Chlorobium thiosulfatophilum.* More than ten years later, in 1987, in the work developed by Arieli *et al*.[36] with the cyanobacterium *Oscillataria limnetica*, it was shown that the adaptation of these bacteria to anoxygenic photosynthesis is based on proton translocation, connected to sulfide-dependent electron transport in thylakoids[36]. The purification and characterization of *Oscillataria limnetica*'s SQR[37] was accomplished only almost 10 years later, in 1994. Research on SQRs expanded after the publication of this characterization.

With the increasing discovery of new SQRs, Theissen and colleagues proposed, in 2003, the first categorization of different SQR types by comparison of common and distinct parts of the available amino-acid sequences.[38] Hence, SQRs were categorized into three groups nonuniformly distributed across eubacteria, archaebacteria, and eukaryotes: group I, containing only eubacterial enzymes; group II, comprising eubacterial and eukaryotic homologs; and group III, including enzymes detected in both eubacteria and archaebacteria.

In 2009, Brito *et. al.* at ITQB NOVA determined the first crystallographic structure of an SQR, that of *Acidianus ambivalens* (PDB: 3H8L)[39,40] . Later that year, a second crystal structure was determined, that of *Aquifex aeolicus* (PDB: 3HYV)[41,42] , followed by the one from *Acidithiobacillus ferroxidans*. Interestingly, with the advent of genome sequencing and more sequences being reported, a new sequence analysis showed that the previous division was not suitable for all new SQR structures.[39]

Therefore, in 2010, Marcia *et al*.[43] proposed a new SQR classification. This new classification had as a base the alignment and comparison of SQR amino acid sequences for which the protein structures were already resolved. In this new classification system, SQRs are divided into six types, I to VI, based on sequence fingerprints which allowed the identification of key residues involved in sulfide metabolism.

1.6.2 Structural insights

The structure of SQR's is comprised of two similar domains, a characteristic shared among the FDR family of enzymes.[39] The structure entails a complex arrangement of a 'twisted five-stranded parallel β-sheet' accompanied by a three-stranded anti-parallel β-sheet on one side and three α-helices on the other side, the so-called Rossman Fold. [3]

The majority of SQR enzymes have in common the presence of two strictly conserved cysteines that form a catalytic disulfide redox center.[3,39,43] In SQR types I, IV, V and VI, a third cysteine residue is also common, which is located at the N-terminal and binds a flavin adenine dinucleotide (FAD) cofactor.

The cysteine sulfur atom is bound to the 8-methylene group of the isoalloxazine ring through a covalent thioether bond.[3,41] An example of one SQR active center with three cysteine residues is represented in Figure 1.8.

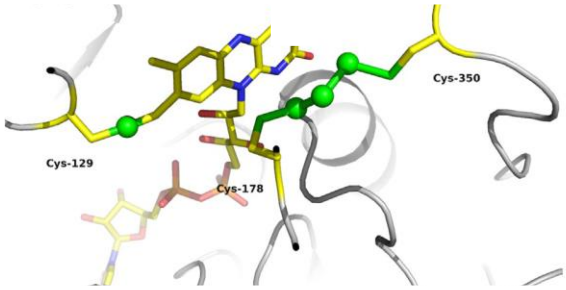

Figure 1.8: **Exemple of a SQR active center showing the three important cysteine residues**. *Adapted from Walsh, et al (2019.*)

In SQR types II and III, as well as in bacterial type V, a distinct mechanism is observed. [3,43]In these proteins, the cysteine that conventionally binds to the FAD cofactor is replaced by a tyrosine (or tryptophan) and a serine (or threonine), potentially resulting in a different binding mechanism.

In fact, in both *S. aureus* and humans, which have type II SQRs, this cysteine residue can be replaced for a tyrosine (or tryptophan) forcing a different mechanism for FAD binding. [3,27,33] In this case, the *S. aureus* SQR enzyme is a type II SQR that contains three cysteine residues. The first residue to account, Cys133, is responsible for binding the FAD cofactor. The conserved cysteine residues, Cys167 and Cys344, play a fundamental role in the active center.

1.6.3 Putative mechanism

From the many possible mechanisms hypothesized in the literature, the formation of a bridging trisulfide is the proposed mechanism confirmed by structural data. Studies performed by Landry *et al* (2019)[44], taking into consideration the bacterial SQR structures obtained by Brito *et al* (2019)[39], guided the proposition of a new mechanism for SQRs way of action (Figure 1.9).

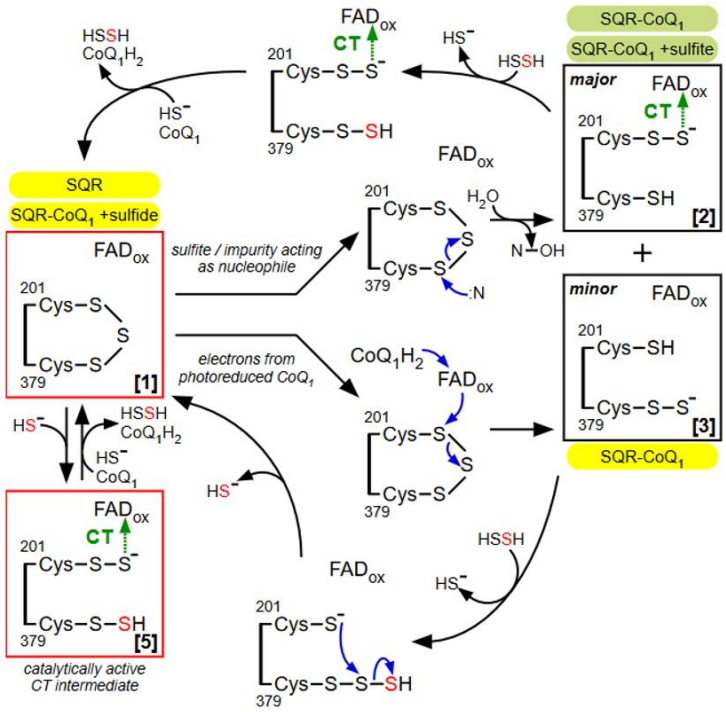

Figure 1.9: **Schematic representation of the mechanism of action for the SQR enzyme.** In the red boxes are presented the active forms of the enzyme ([1] and [5]) showing the presence of a trisulfide. This conformation Is lost leading to the formation of two non-reactive sulfide species (black boxes): [2] ^{201}Cys-S$_\gamma$-S$^-$ major species and [3] ^{379}Cys-S$_\gamma$-S$^-$ minor species. The trisulfide is rebuilt in the presence of HSSH through one of the two mechanisms presented ([3] to [1] or [2] to [1]). *Adapted from Landry, et al (2019)* [44]

The proposed mechanism suggests that the resting state of SQR contains a cysteine trisulfide cofactor (Cys_S$_\gamma$-S-S$_\gamma$_Cys).[44] However, when the presence of CoQ1 leads to the disruption of this trisulfide bond and to the formation of two different sulfur species: ^{201}Cys_Sγ-S$^-$ and ^{379}Cys_S$_\gamma$-S$^-$. The exact source of this disruption is not clear, but it could be due to reduction or nucleophilic addition. This implies that the enzyme's active state involves this intact trisulfide bond.[44]

The presence of HSSH, helps rebuilding the trisulfide bridge in one of two ways. On one hand, the attack on the S of the exposed Cys^{379} in the resting trisulfide releases ^{201}Cys-S$_\gamma$-S$^-$ that binds with FAD to form the charge transfer (CT) complexes. With reduction of CoQ1 the active form of the enzyme is reconstituted. On other hand, the presence of HSSH can lead to the restoration of the active form of SQR.

In conclusion, the proposed mechanism involves the unique retention of sulfane sulfur within the active site of SQR, with specific cysteine residues playing crucial roles in sulfur transfer reactions. This mechanism allows organisms to detoxify hydrogen sulfide and use the resulting energy to drive metabolic processes.

1.7 Aim of the project

The major aim of this project is to characterize biochemically and structurally the SQR protein from *S. aureus* (*Sa*SQR). To achieve this goal, a protocol “from gene to structure” was developed for this protein.

Plasmids encoding *Sa*SQR were previously prepared in Giedroc’s Lab (Indiana University, USA). From here, a protocol to express and purify this protein were developed and optimized, in order to achieve high levels of very pure protein for the subsequent crystallization trials. Next, crystallization conditions will be screened and optimized, to produce protein crystals suitable for X-ray diffraction experiments.

Furthermore, the protocols optimized for the wild-type protein will be applied to two mutant proteins: *Sa*SQR_C133S, where the residue Cys133 responsible for the binding of the FAD group is replaced by a serine residue, and *Sa*SQR_C344S, where the Cys344 residue present in the active site is replaced by a serine. The aim is also to perform the biochemical and structural characterization of these mutant proteins.

2

MATERIALS AND METHODS

This project results from a collaboration between the Laboratory of Prof. Dr. David P. Giedroc at the Chemical and Physical Biology of Infectious Disease laboratory, from Indiana University (USA)[45], and the Membrane Protein Crystallography Laboratory, at ITQB NOVA (Portugal)[46]. The project is included in a long-lasting collaboration between Professor Giedroc and Dr. José Artur Brito, who supervises a team involved in Hydrogen Sulfide Homeostasis in Pathogenic Bacteria.

In previous work of the lab, the plasmids of interest, kindly provided by Professor Giedroc, were transformed into different bacterial strains and isolated for storage.

All stored plasmids were sequenced (GATC Biotech - Eurofins[47]) and it was determined that *Escherichia coli* BL21 pRARE was the most suitable strain to express the protein of interest. Using a NanoDrop™ OneC UV-Vis spectrophotometer (Thermo Scientific™), the quantity, quality and integrity of the plasmid isolated from this strain was evaluated before use. The plasmid was stored at -20 ºC at a final concentration of 323.3 ng/mL.

2.1 Buffers composition

During this project several buffers were used, with different compositions, listed in Table 2-1.

Table 2-1: **List of buffers used during this project.**

Buffer	Composition
Buffer L0 (lysis buffer)	50 mM Tris pH 8.0, 500 mM NaCl, 1 mM TCEP, 5 mM $MgCl_2$, 20 µM FAD, 1 mg/mL lysozyme, 0.25 U/mL Benzonase, protease inhibitors (1 pill/100 mL)
Buffer A0 (binding buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 10% (v/v) glycerol, 10 mM TCEP, 20 mM imidazole, 20 µM FAD
Buffer E0 (elution buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 10% (v/v) glycerol, 10 mM TCEP, 500 mM imidazole, 20 µM FAD

Buffer L1 (lysis buffer)	50 mM Tris pH 8.0, 500 mM NaCl, 1 mM TCEP, 5 mM $MgCl_2$, 20 µM FAD, 1mg/mL lysozyme, 0.25 U/mL Benzonase, protease inhibitors (1 pill/100 mL)
Buffer A1 (binding buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 20 µM FAD, 10% (v/v) glycerol, 20 mM imidazole
Buffer B1 (washing buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 20 µM FAD, 10% (v/v) glycerol, 50 mM imidazole
Buffer C1 (washing buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 20 µM FAD, 10% (v/v) glycerol, 75 mM imidazole
Buffer D1 (elution buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 20 µM FAD, 10% (v/v) glycerol, 500 mM imidazole
Buffer S1 (storage buffer)	50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 20 µM FAD, 10% (v/v) glycerol
Buffer L (lysis buffer)	50 mM Tris pH 8.0, 500 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, 20 µM FAD, 1mg/mL lysozyme, 0.25 U/mL Benzonase, protease inhibitors (1 pill/100 mL)
Buffer A (binding buffer)	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 20 mM imidazole
Buffer B (washing buffer)	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 50 mM imidazole
Buffer C (washing buffer)	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 75 mM imidazole
Buffer D (elution buffer)	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 500 mM imidazole
Buffer S (storage buffer)	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, 20 µM FAD, 10% (v/v) glycerol
Buffer L2 (lysis buffer)	50 mM succinic acid pH 6.6, 500 mM NaCl, 0.2 mM DTT, 5 mM $MgCl_2$, 20 µM FAD, 1mg/mL lysozyme, 0.25 U/mL Benzonase, protease inhibitors (1 pill/100 mL)
Buffer A2 (binding buffer)	50 mM succinic acid pH 6.6, 150 mM NaCl, 0.2 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 20 mM imidazole
Buffer B2 (washing buffer)	50 mM succinic acid pH 6.6, 150 mM NaCl, 0.2 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 50 mM imidazole
Buffer C2 (washing buffer)	50 mM succinic acid pH 6.6, 150 mM NaCl, 0.2 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 75 mM imidazole
Buffer D2 (elution buffer)	50 mM succinic acid pH 6.6, 150 mM NaCl, 0.2 mM DTT, 20 µM FAD, 10% (v/v) glycerol, 500 mM imidazole
Buffer S2 (storage buffer)	20 mM succinic acid pH 6.6, 150 mM NaCl, 0.2 mM DTT, 20 µM FAD, 10% (v/v) glycerol

2.2 Transformation of competent cells

E. coli XL10-Gold ultracompetent cells were designed for cloning large plasmids and ligated DNA with the highest transformation efficiency possible, while exhibiting fast growth and large colonies.[27] For these reasons, the SQR expression plasmid (*sqr* cloned into a pHis parallel expression plasmid[48] -Figure 2.1) was first transformed into XL-10 gold competent cells.

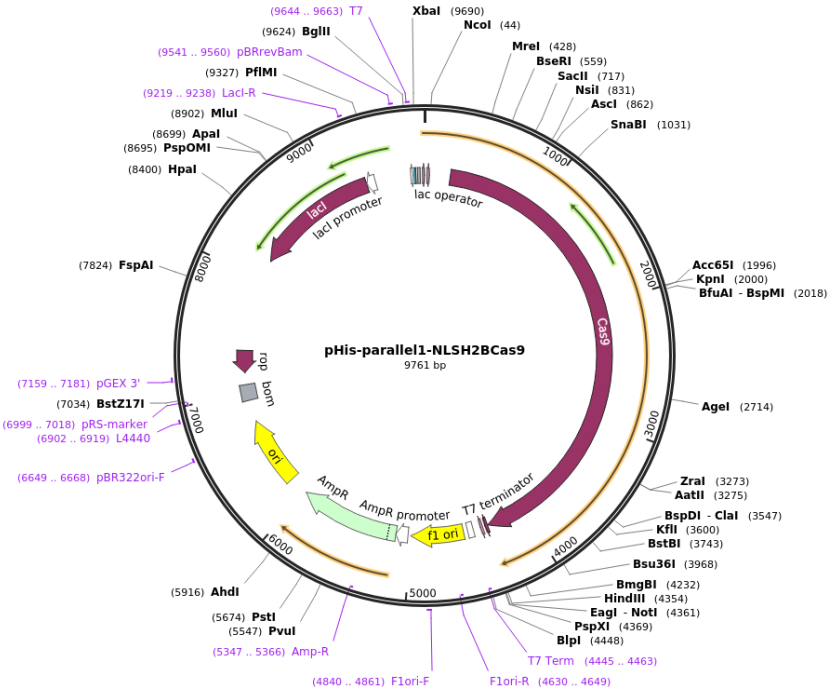

Figure 2.1: **Sequence map of pHis-parallel1-NLSH2BCas9 plasmid**.

Afterwards, the plasmid was isolated from these cells and transformed into E. *coli* ArcticExpress cells[49], which are derived from the high-performance *E. coli* BL21-Gold competent cells and engineered to address the common bacterial gene expression hurdle of protein insolubility. These cells enable efficient high-level expression of heterologous proteins in *E. coli*.

Glycerol stocks of the competent cells, stored at -80 ºC, were thawed on ice and incubated with 100 ng of plasmid for 30 minutes. The mix was incubated in a water bath at 42 ºC, for one minute, and immediately transferred to ice for two minutes. The cells were allowed to recover in 1 mL of Luria-Bertani (LB) medium and incubated at 37 ºC, with orbital shaking at 220 rpm, for 1 hour. The cells were collected at 5000 rpm, for 2 minutes, at room temperature or 4 ºC, and resuspended in 200 µL of supernatant.

2.3 Overexpression of SQR proteins

The growth conditions for the cell cultures were established according to the work of Shen *et al.*[4] and previous work performed in the laboratory.

Starter cultures were prepared by inoculating 50 mL of LB, supplemented with the appropriate selection antibiotic. *E. coli* XL-10 Gold cells were selected in LB-agar supplemented with 100 µL/mL ampicillin and *E. coli* ArticExpress cells were selected in LB-agar supplemented with 10 µL/mL gentamycin, with a single colony, followed by incubation at 37 ºC and 220 rpm, for 16 hours (MaxQ 4000 E Class Orbital Incubator Shaker, Barnstead Lab Line).

These cultures were diluted to 1.5% (v/v) in 1 L of fresh LB supplemented with 100 mg/mL ampicillin and 10 mg/mL gentamycin, and incubated at 37 ºC and 220 rpm (SH Floor R, Controltecnica) until reaching an optical density at 600 nm (OD_{600}) between 0.6 and 0.8.[27]

The sequence that encodes for the protein of interest (see section A.2.1) was insert in a pHis plasmid. The construct is resistant to ampicillin, resistance acquired from the plasmid, and it is resistant to gentamicin, resistance acquired by the insert. The insert has a 6xHistag that will allow its identification in western blot and the purification of the protein due to the affinity of histidine residues to the nickel column. The plasmid used has a *lac* operon that becomes transcriptionally active when lactose is present, and is repressed in its absence. In this experiment, isopropyl β-D-1-thiogalactopyranoside (IPTG), a non-metabolizable lactose analogue was used at 1 mM.[50]

The cells were further incubated for 6 hours at 15 ºC and 220 rpm. At the end, cells were collected at 8000 rpm and 4 ºC for 10 minutes, and stored at -80 ºC in 75% (v/v) glycerol.

2.4 Cell disruption for protein isolation

According to previous work developed in the Lab, the preferential disruption method for the protein of interest is the mechanical method of sonication. For this, frozen cell pellets were thawed and resuspended in **Buffer L**. The cells were disrupted by sonication (UP200S Ultrasonic Processor, Hielscher) with 8 cycles of 30 seconds, 50 % amplitude, 0.6 s/pulse, each followed by a 30 seconds cooling interval on ice. The cells suspension was centrifuged at 8400 rpm, for 10 minutes, at 4 ºC, (benchtop centrifuge 5804-R, Eppendorf) and the supernatant was kept on ice until purification.

2.5 Protein purification

As a first purification approach, the clarified soluble fraction was injected in a high-performance immobilized metal affinity chromatography (IMAC) column, for His-tag recombinant protein purification. Since this approach was not efficient, purification through gravity-flow using a HisGraviTrap column was tested (Figure 2.1-Approach 1).

Firstly, the column was equilibrated with 10 mL of Buffer A1. After applying the sample, the column was washed with Buffer C1, and elution was achieved with Buffer D1 (Figure 2.1-Approach 2).

After optimizing buffer composition using Thermal Shift Assay and NanoDSF (see sections 2.7 and 2.8) data, the column was equilibrated with 10 column volumes (CV) of **Buffer A**. After sample application, the column was washed with 10 CV of **Buffer B** and **Buffer C,** containing 50 mM and 75 mM imidazole, respectively (Figure 2.1-Approach 3).

The elution was executed using 3.5 CV of **Buffer D**, with a final concentration of 500 mM imidazole. The eluted protein was collected in fractions of 0.3, 0.7 and 2.5 mL. Imidazole was removed from the sample, as it can interfere with the crystallization process, using a PD10 desalting column, as described by the manufacturer. Briefly, the column was equilibrated with **Buffer S,** an imidazole-free buffer and the sample was collected in the same buffer.

The sample was concentrated using a 30 kDa MWCO MilliporeAmicon® Ultra Centrifugal Filter Unit, before injection in a Superdex 200 Increase 10/300 GL column equilibrated with **Buffer S**, as a final polishing step. This column allows high resolution gel filtration in a small scale (mg) preparative purification. This step was later on eliminated since, it did not significantly improve the purification process.

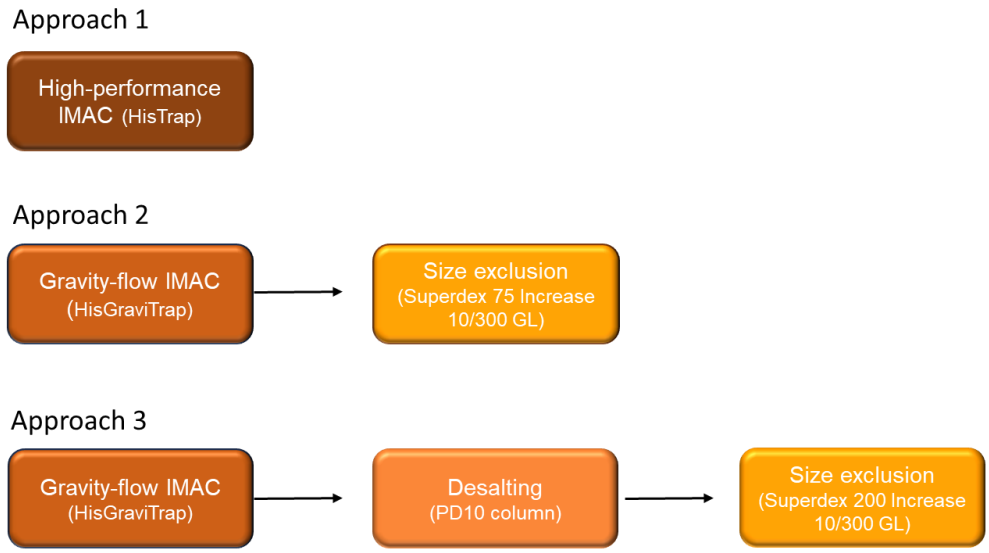

Figure 2.2: Schematic representation of the different approaches adopted for the optimization of the purification protocol for *Sa*SQR.

SDS-PAGE was performed in 12 % (w/v) polyacrylamide gels, at 180 V, using MOPS as running buffer. Samples collected during purification were denatured at 95 ºC, for 10 minutes, in the presence of β-mercaptoethanol. The gels were stained with Coomassie Brilliant Blue (ProBlue Safe Stain, Giotto BioTech, Italy) and imaged using a Gel Doc™ EZ system (Bio Rad).

2.6 Western Blot

The western blot was performed with the wild type and both mutants. It guaranteed the detection and identification of the proteins in the different fractions by the presence of an 6xHistag, as mention before in this chapter. The purified proteins were loaded into a 12 % (w/v) polyacrylamide gel (SDS-PAGE) and run at 180 V, for 45 minutes. The gel was then washed in 1X Transfer Buffer (TBS) with methanol during approximately 5 minutes. After this time the gel was transfered to a nitrocellulose membrane equilibrated in 1X TBS using a Mini Blot Module (Thermo Fisher Scientific), with the Blot StandardSD program (1.0 A, 25 V, 30 minutes).

When finished, the membrane was washed in 20 mL of TBS mixtured with polysorbate 20 (TBST) for 5 minutes. The membrane was incubated in a blocking solution (5 % (w/v) non-fat milk diluted in 20 mL 1x TBST), for 1 hour with constant shaking, and then washed 3 times with 1X TBST. The primary antibody solution (HisProbe-HRP diluted 1: 5000 in 1X TBST) was applied over the membrane for 1 hour at room temperature with gentle agitation and the membrane was again washed three times for 5 minutes with 20 mL 1X TBST.

Chemiluminescent detection was performed using the Supersignal west pico chemiluminescent substrate reagent following the manufacture protocol. Signals were visualized and photographed with ChemiDocTM XRS+ System (Bio-Rad).

2.7 Thermal Shift Assay ("Thermofluor")

As an initial step to determine the ideal amount of protein and dye to use for this assay, different amounts of protein (1, 2, 5 and 10 µg) were added to different concentrations of SYPRO Orange dye (1, 2, 5 and 10x). The best condition determined use 5 µg of protein and 2x dye concentration. These conditions were used for the remaining of this work.

Thermal shift assays were employed to optimize the protein buffer regarding protein stability using the commercially available kits Durham pH Screen and RUBIC additive screen (both from Molecular Dimensions, UK) (The compositions of these screens are listed in appendix A.3 and A.4).

To perform thermal denaturation, a QuantStudio 7 Flex Real-Time PCR system (*Applied Biosystems*™), was used heating the samples in a range of temperature between 20°C to 95°C, with a temperature gradient of 1°C/min, resulting on a denaturation curve.

2.8 Nano differential scanning fluorimetry

Nano Differential Scanning Fluorimetry (NanoDSF) assays started with a first discovery run performed to determine the most suitable excitation percentage for the assay. In this run, the capillary type (normal or high sensitivity) was also tested. For these tests, approximately 2.5 µg of protein were used per condition, in a final volume of 10 µL.

To test different buffer compositions, 2.5 µg of protein in Buffer S were used per condition, in a final volume of 15 µL containing the additive/mixture of additives to be tested: 5 mM DTT, 5 mM Octyl-β -Glucoside Detergent (OG), 5 mM biotin, 50 mM glucose, 50 mM arginine, or 50 mM glutamate. To allow comparison between the different conditions, reference runs were performed with only water and **Buffer S**.

2.9 Crystallization screens

From the first purification approach, an initial crystallization test was performed with a protein concentration of 9.4 mg/mL in Buffer S, using the sitting drop technique. The drops were prepared using 100 nL of protein and 100 nL of precipitant, equilibrated against 40 µL of reservoir solution, in 96 wells TPP plates (LabTech) at 20 ºC. The BCS (Basic Chemical Space) and PACT (pH, Anion, Cation crystallization Trial) Premier crystallization screens (both from Molecular Dimensions, UK) were used as precipitants.

After performing a Pre-Crystallization Test (PCT, Hampton Research, US), further crystallization trials were carried out with 1.5 and 3.0 mg/mL of purified protein in Buffer S, resulting from the second purification approach. The sitting drop technique was prepared similarly as before but using SwisSci 96-Well 2-Drop Plates, incubated at 4 ºC and 20 ºC.

3
RESULTS AND DISCUSSION

In this section, the results of this project are presented and divided into three primary sections. The first section comprises the optimization of *Sa*SQR expression, including evaluation of plasmids, bacterial expression strains, and culture conditions (culture media, temperature, time, etc.). The second section outlies the efforts undertaken to develop a purification protocol for *Sa*SQR, in order to obtain suitable amounts of protein for crystallization trials. Lastly, the third section describes the several crystallization trials that were performed, along with additional thermostability assays to determine the most favorable buffer conditions for stabilization of the target protein.

3.1. Assessment of *Sa*SQR constructs integrity and expression efficiency

The protein under study, *S. aureus*' sulfide:quinone oxidoreductase, from now on referred to as *Sa*SQR_WT or just WT, is a membrane-associated protein that has 415 amino acid residues. In this project, this protein was expressed with a N-terminal 6xHis-tag to aid in its purification. The total construct has a theoretical molecular weight of 47.8 kDa, as predicted by ProtParam from the ExPASy database (see section A.1).[51]

*Sa*SQR has three amino acid residues of particular interest: the Cys167-Cys344 pair, proposed to form redox active center, and Cys133, the residue that binds the FAD cofactor. Previous work in the Giedroc Lab produced two mutant variants of *Sa*SQR: a Cys133Ser (*Sa*SQR_C133S), and a Cys344Ser (*Sa*SQR_C344S) variant. The plasmids designed to express these genes were transformed into three different *E. coli* bacterial strains, chosen according to *Shen et al.* (Giedroc's Lab): Rosetta, Rosetta Gami2 and BL21 pRARE. The plasmids were isolated from these strains using the NZYMiniprep commercial kit and stored at -20 ºC. An additional WT plasmid was stored amongst these plasmids, which was tested and yielded favorable preliminary expression results.

Therefore, to evaluate the integrity of the plasmids isolated and stored in previous works, and to determine which ones to use, all plasmids were digested with the restriction enzyme *XbaI*. This enzyme was selected because it cuts the plasmid in only one site and is similar to the pRARE plasmid.

This plasmidhas been demonstrated to improve expression yields in E. *coli* strains by facilitating the expression of tRNAs that correspond to non-common codons important for protein synthesis in this bacterial system.The digestion products were analyzed in a 1% (w/v) agarose gel (Figure 3.1).

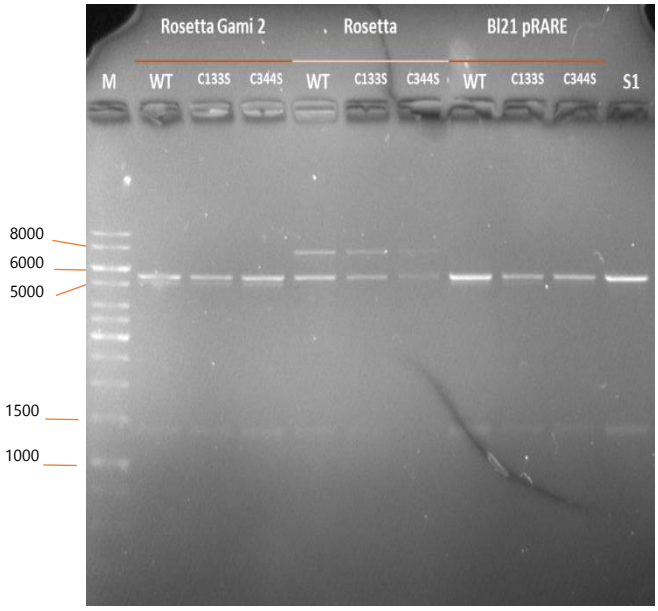

Figure 3.1: ***XbaI* digestion of pHis plasmids with SQR inserts**. The plasmids (1 µg) and the restriction enzyme were incubated together for 1 hour. The enzyme used has a single restriction site in both the plasmid of interest and pRARE. Digestion products were loaded on a 1% agarose gel, run at 70 mV. M: Generuler 1kb DNA Ladder marker[51]; WT: *Sa*SQR wild type; C133S, C344S: *Sa*SQR mutants; S1: *Sa*SQR WT used successfully in previous works.

Two bands were detected in all samples: one of approximately 1 500 bp, corresponding to the *Sa*SQR insert, and another of approximately 5 000 bp, corresponding to the pHis plasmid. For the Rosetta strains, an additional unexpected band was observed around 6 000 bp, and the insert band was the faintest out of the three strains tested. For these reasons, plasmids isolated from Rosetta cells were not used. Rosetta Gami2 and BL21 pRARE show a double band around 5 000 bp, which corresponds to the pHis and pRARE plasmids, both of similar size. The BL21 pRARE plasmids were selected to be used in this project as they were the most similar to well S1, *i.e.*, the most similar to the plasmid that previously yielded good expression results. To further assess the BL21 pRARE plasmids, these were sequenced to confirm the integrity of the inserts. Sequence alignments with the His-tagged *Sa*SQR wild-type and mutant proteins is showen in section A.2.

Competency relates to the ability of a cell to uptake foreign DNA. It can occur naturally, or it can be induced by permeabilization of the membrane. In this case, cells were treated with $CaCl_2$, as the Ca^{2+} ions promote DNA binding to the cell membrane thus improving its stability. The heat shock method was used for the transformation (incorporation of exogenous DNA) of the *E. coli* Arctic express competent

cells. This method causes an alteration of the membrane fluidity and pore size, which facilitates the entrance and entrapment of exogenous DNA.

After the transformation (see section 2.2), cells were cultured for 16 hours in LB medium supplemented with the appropriate selection antibiotics. Cell lysis was performed as described in *Shen et al*. 2016, and the resulting soluble and insoluble fractions were analyzed by SDS-PAGE. A western blot analysis (section 2.6) was also performed to confirm the identity of the protein bands.

From the SDS-PAGE and western blot analyses (Figure 3.2), it was possible to conclude that there was protein expression in all samples tested. Better expression profiles were obtained for the mutant proteins compared to the WT. Regarding the WT protein, after lysis, the protein was present in both the soluble and insoluble fraction. As for the mutants, following cell lysis most of the protein was in the insoluble fraction.

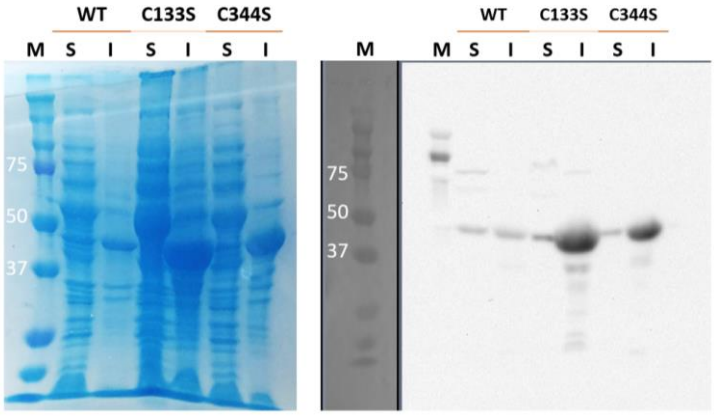

Figure 3.2: **Evaluation of the efficiency of *Sa*SQR expression and cell lysis using BugBuster. Right panel:** the soluble (S) and insoluble (I) fractions resulting from cell lysis were analyzed by SDS-PAGE, performed on a 12% polyacrylamide gel at 180 V. M: precision plus protein dual color standards (Biorad) . **Left panel:** analysis of the same samples by Western Blot, using 1:5000 HisProbe-HRP (Sigma).

Furthermore, it was also possible to confirm that the proteins migrate in the gel in accordance with their theoretical molecular weight (47.8 kDa, as calculated by ExPASy).

The ultimate goal was to establish expression and purification protocols that were reproducible and yield high amounts of WT protein, allowing further studies on this protein to be done. For this reason, plasmids encoding mutant *Sa*SQRs were stored at -20 ºC, and subsequent tests were only performed with cells transformed with plasmids encoding the WT protein. Since the expression protocol did not work as expected in Shen *et al* [27], further expression tests were undertaken.

Cells were cultured in TB medium for 6 hours[8], up to an OD_{600} between 0.8 and 1.2, and protein expression was induced with 0.25 mM of IPTG. Culture aliquots were collected after 1, 2, 4, 6 and 16 hours of induction, and cells were chemically lysed using BugBuster. After centrifugation, the soluble and insoluble fractions of each sample were analyzed by SDS-PAGE (Figure 3.3).

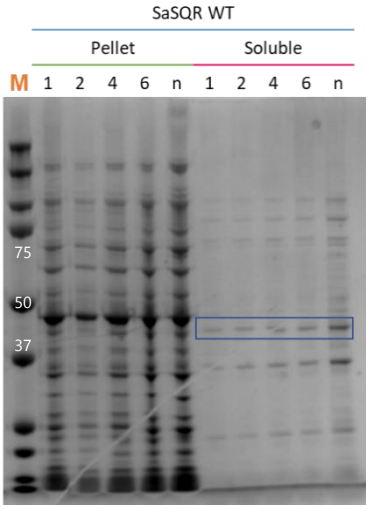

Figure 3.3: ***Sa*SQR expression tests**. *E. coli* Artic Express cells were cultured in TB and *Sa*SQR expression was induced with 0.25 mM IPTG, after 6 h of growth. Samples were collected at 1, 2, 4, 6, and 16 (n) h post induction, and lysis was achieved using BugBuster. The soluble and insoluble fractions of each sample were analyzed by SDS-PAGE, in 12% polyacrylamide gels run at 180 V. M: precision plus protein dual color standards (Biorad).

By inducing protein expression at a later growth point, the levels of protein expression increased. From the results obtained in this test, it was possible to conclude that 16 hours of incubation result in higher protein yields, although 6 hours also result in good expression levels. Protein expression depends on multiple external factors, such as, nutrient availability, toxicity of the media, pH, among others. Unbalance on this factors can lead to a misexpression of protein, lost of its stability or other problems connected to expression conditions. As the protocol used by the person working on this project previously included 16 hours of induction, to discard the possibility that prolonged expression times lead to loss of protein stability, we chose to do 6 hours of induction and assessed the results. This proved to be a successful change and all the results presented in the next chapter were obtained with 6 hours induction.

3.2. Protein purification

The first *Sa*SQR_WT purification attempt was based on previous work in the lab. Since *Sa*SQR was expressed with an His-tag, and histidine residues have high affinity for nickel, it was possible to isolate this protein from the crude cell lysate using a nickel column. In affinity chromatography the target protein is specifically and reversibly bound to the column resin. The sample was applied under conditions that favor specific binding to the ligand. Unbound material was washed out of the column. Later, the bound target protein was recovered by changing the conditions to favor elution. Elution can be performed specifically, using a competitive ligand, or nonspecifically, by changing for example the pH, ionic strength, or polarity. In this case, elution was promoted by competition using a buffer with high concentrations of imidazole.

Prior to protein purification, cell pellets were thawed, homogenized in lysis buffer, and disrupted by sonication. Cells were disrupted according to the procedure described in section 2.4 using **Buffer L0**. The clarified cell lysate was injected in a 5 mL HisTrap column equilibrated in **Buffer A0** (containing 20 mM imidazole), connected to a pressurized system (ÄKTA pure). According to previous results, the protein of interest was eluted at around 60% of **Buffer E0** (500 mM imidazole), while the contaminants elute around 30%. For this reason, a stepwise elution was performed with steps of 0, 30, 70 and 100% of **Buffer E0** (section 2.1).

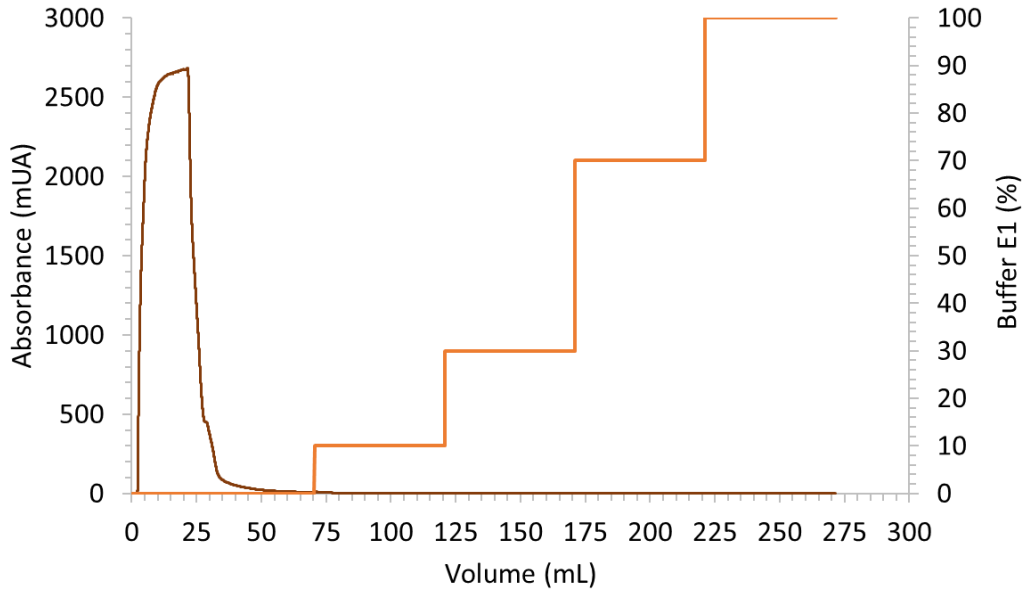

Figure 3.4: **Purification of *Sa*SQR by affinity chromatography.** Elution profile of *Sa*SQR using a HisTrap column on a ÄKTA Pure system. The protein of interest was eluted by an imidazole stepwise gradient, from 20 to 500 mM. Absorbance at 280 nm is represented in brown, and the percentage of elution buffer (E0) is represented in orange.

As it can be observed in the chromatogram (Figure 3.4), this purification step was not successful as no protein peaks were detected. Among other possible justifications for this result, two determining factors could have been that the column was not in proper condition for use, or that the protein loses stability when in pressurized environments.

Therefore, purification by affinity chromatography was pursued in a non-pressurized system. For this, gravity-flow HisGraviTrap columns were used. Purification was pursued using **Buffers A1**, **C1** and **D1**. Since this set-up does not allow absorbance tracking like on ÄKTA systems, samples were collected during elution and analyzed by SDS-PAGE (Figure 3.5).

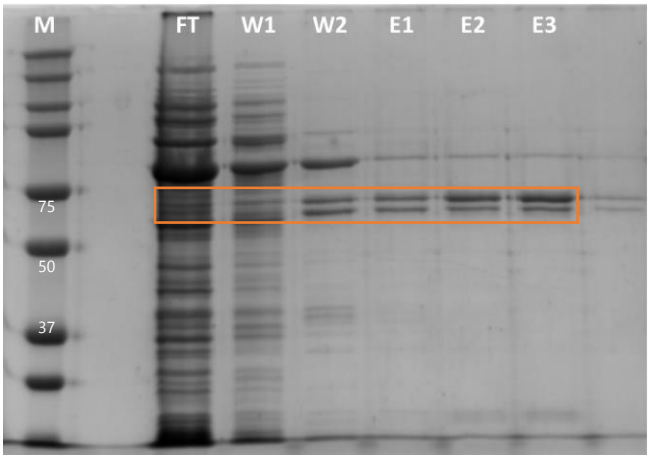

Figure 3.5: **SDS-PAGE analysis of samples collected during *Sa*SQR purification by affinity chromatography.** *Sa*SQR was purified using a gravity-flow HisGraviTrap column. Samples were collected during elution and analyzed by SDS-PAGE, in 12% polyacrylamide gels run at 180 V. M: precision plus protein dual color standards (Biorad); FT: sample application flow-through; W1: first column wash (20 mM imidazole); W2: second column wash (75 mM imidazole); E1, E2, E3: elution with 500 mM imidazole.

The two column wash steps were successful in removing most of the contaminants present in the crude cell lysate. However, the gel lanes corresponding to protein elution show a double band profile(highlighted with the orange box): an additional band of over 50 kDa was present alongside the protein of interest, possibly being a protein from the *E. coli* host. To further purify *Sa*SQR and achieve a final homogenized sample with a single oligomeric state, the fraction collected from the second elution was concentrated down to 200 µL and injected in a size exclusion column (Figure 3.6), with a 200 µL capillary loop.

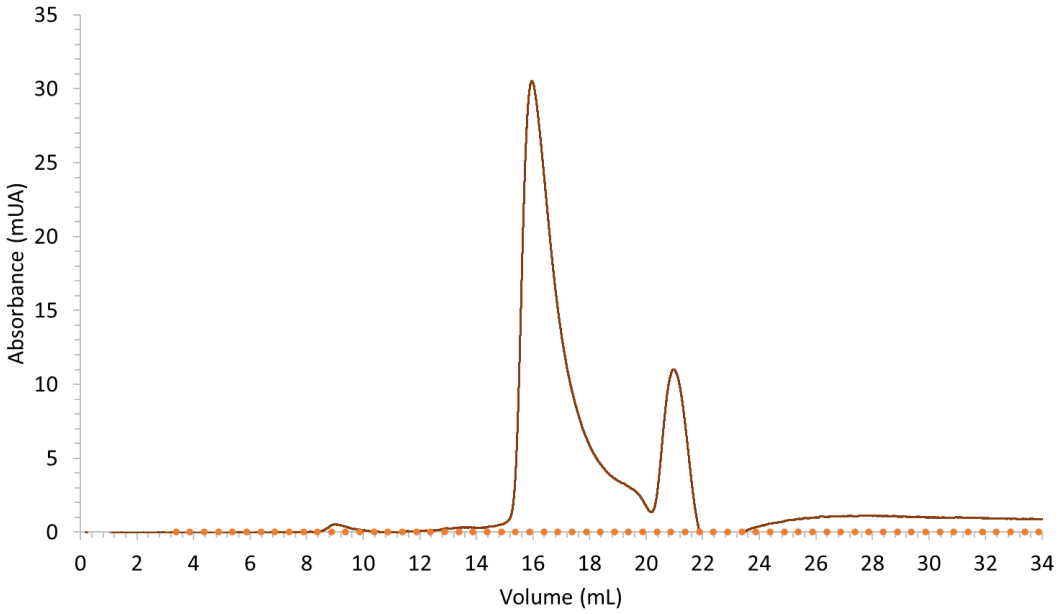

Figure 3.6: **Purification of *Sa*SQR by size exclusion chromatography**. Elution profile of *Sa*SQR using a Superdex 75 Increase 10/300 GL column on a ÄKTA Pure system. The protein injected in the column was in Buffer D1 and elution was performed with buffer S1. Absorbance at 280 nm is represented in brown, and the fractions collected are represented by orange dots.

It is possible to see a peak between 15 and 20 mL of elution and another peak around 20-23 mL. Fractions corresponding to this peak were analyzed by SDS PAGE (Figure 3.7).

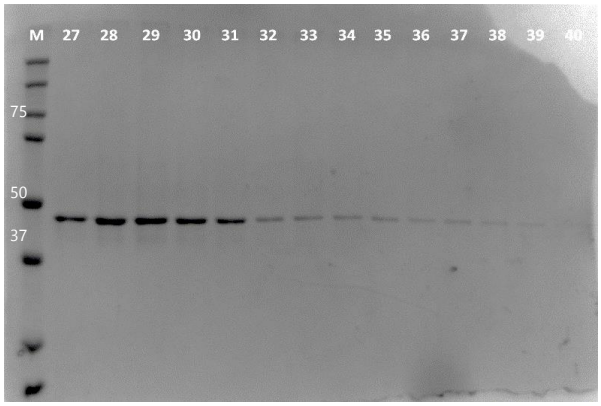

Figure 3.7: **SDS-PAGE analysis of the efficiency of the SEC purification**. Following affinity purification, *Sa*SQR was purified by SEC using a Superdex 75 column equilibrated with buffer S1. Samples were collected during elution and analyzed by SDS-PAGE, in 12% (w/v) polyacrylamide gels run at 180 V. M: precision plus protein dual color standards (Biorad). The number indicated in each well corresponds to the number of the fraction collected.

These results allow us to understand that the first peak corresponds to *Sa*SQR, the protein of interest. Since the size exclusion column separates molecules according to their size, bigger molecules will elute before than smaller molecules, appearing first in the chromatogram, the second peak might correspond to a different oligomeric form of *Sa*SQR. This hypothesis is supported by the fact that, in the SDS-PAGE (Figure 3.7) all the fractions have a band with molecular weight corresponding to the protein of interest.

Both in the chromatogram and in the SDS-PAGE it is possible to observe that the amount of protein is not very high. In the chromatogram the absorbance levels were low and in the gel the protein bands, despite being visible and well defined, were not very intense.

With this first trial, we concluded that the washing steps of the HisTrap affinity purification can be improved, to remove more contaminants from the crude cell lysate. Moreover, it was possible to retrieve the current elution profile of the protein.

Having this in mind, a new purification approach was tested using **Buffers A1**, **B1**, **C1** and **D1**, according to the protocol described in section 2.5, using a HisGraviTrap column. The fractions that had higher levels of the target protein, and fewer contaminants, were pooled together and applied onto a PD10 column for buffer exchange, from **Buffer D1** (500 mM imidazole) to **Buffer S1** (imidazole-free). This allowed the removal of the imidazole present in the previous buffers, which can influence protein stability and crystallization. All fractions were analyzed by SDS-PAGE (Figure 3.8).

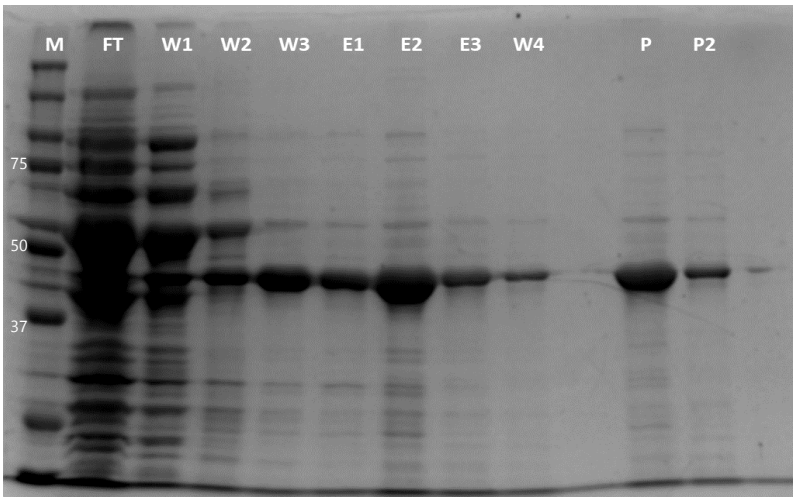

Figure 3.8: **SDS-PAGE analysis of the efficiency of the affinity chromatography purification**. *Sa*SQR was purified using a gravity-flow HisGraviTrap column. Samples were collected during elution and were analyzed using a 12% poly-acrylamide run at 180 V. M: precision plus protein dual color standards (Biorad); FT: sample application flow-through; W1: first column wash (20 mM imidazole); W2: second column wash (50 mM imidazole); W3: third column wash (75 mM imidazole); E1, E2, E3: elution with 500 mM imidazole; P: E2 fraction after desalting; P2: E1+E3+W4 fractions after desalting.

It is possible to see that the addition of an extra washing step results in a cleaner protein elution (Figure 3.8., lane E1). By comparing lanes E2 and P, i.e., *Sa*SQR before and after elution from the PD10 column, we can confirm that the protein profile remains unaltered after buffer exchange, with minimal dilution. Altogether, these constitute improvements to the previous protocol for the first stage of protein purification.

Following buffer exchange, the sample was injected in a Superdex-200 column (Figure 3.9). This time, a different column was used to ensure a better separation between molecules with similar molecular weights.

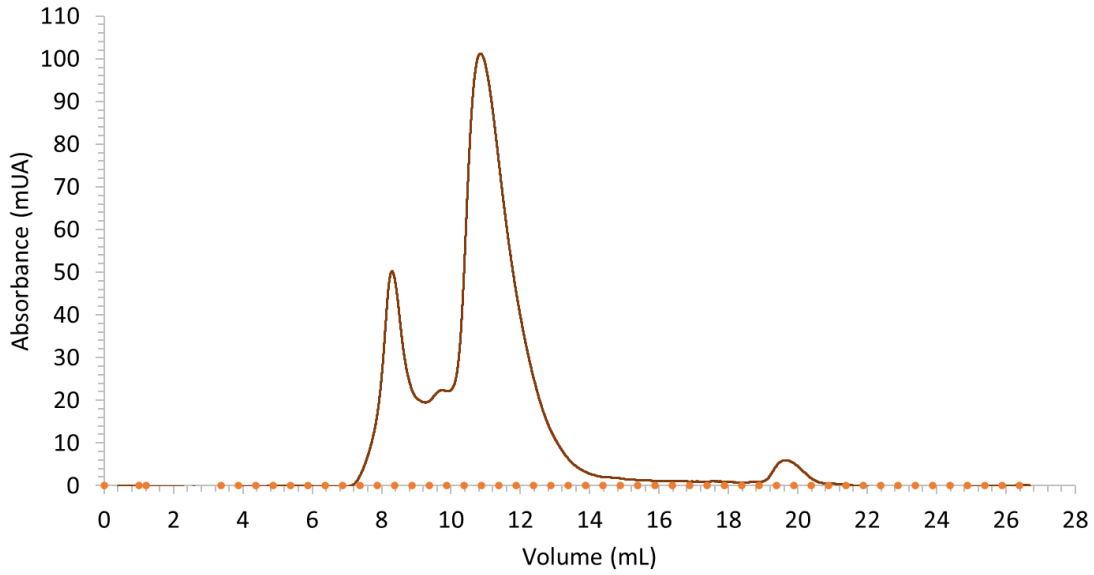

Figure 3.9: **Purification of *Sa*SQR by size exclusion chromatography.** Elution profile of *Sa*SQR using a Superdex 200 Increase 10/300 GL column on a ÄKTA Pure system. The protein of interest was injected in buffer S (Imidazole free) and elution was performed in the same buffer. Absorbance at 280 nm is represented in brown, and the fractions collected are represented in orange.

Similarly to what happened in the first trial, it was possible to distinguish three elution peaks between 7 and 9 mL (fractions 12-15), 10 and 14 mL (fractions 17-24), and 19 and 21 mL (fractions 35-38). Some of these fractions were analyzed by SDS-PAGE (Figure 3.10).

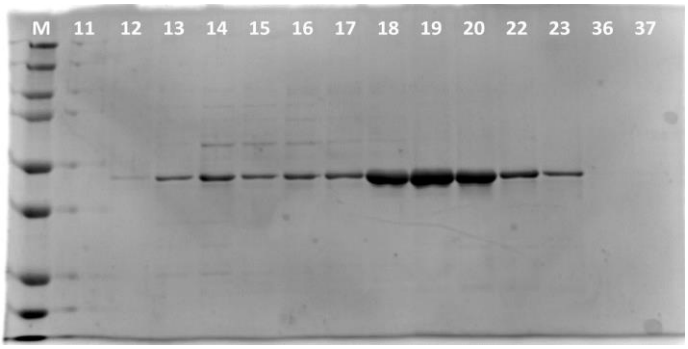

Figure 3.10: **SDS-PAGE analysis of the efficiency of the SEC purification**. Following affinity purification, *Sa*SQR was purified by SEC using a Superdex 200 column equilibrated with buffer S1. Samples were collected during elution and analyzed by SDS-PAGE, in 12% polyacrylamide gels run at 180 V. M: precision plus protein dual color standards (Biorad). The number indicated in each well corresponds to the number of the fraction collected.

Once again, the larger peak (fractions 15 -24) corresponds to *Sa*SQR, in a very pure state. These fractions were pooled together and concentrated to 3.6 mg/mL. Part of this purified protein was used in the first crystallization screen, while another was used for nanoDSF and TSA assays (section 3.3).

From these assays, it was initially thought that the buffer best suited for *Sa*SQR was succinic acid pH 6.6. Therefore, buffers with succinic acid were prepared (Buffer A2, B2, C2, D2 and S2) and used in the protocol for protein purification. Although the HisGraviTrap purification step had similar results to the ones obtained using Tris pH 8.0, no protein peaks were detected in the SEC step. Using the data described in section 3.3, a new set of buffers was designed.

To test the new set of buffers and ensure that the protocol was reproducible, the purification protocol was repeated from the beginning: purification by HisGraviTrap with three washing steps, using **Buffers A, B, C** and **D** (Figure 3.11), followed by removal of imidazole using **Buffer S** and finally SEC purification using a Superdex 200 Increase column (Figure 3.12 and Figure 3.13).

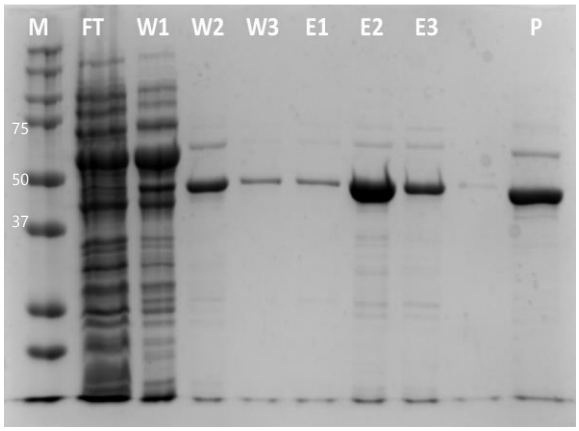

Figure 3.11: **SDS-PAGE analysis of the efficiency of the affinity chromatography purification**. *Sa*SQR was purified using a gravity-flow HisGraviTrap column. Samples were collected during elution and were analyzed using a 12% (w/v) polyacrylamide run at 180 V. M: precision plus protein dual color standards (Biorad); FT: sample application flow-through; W1: first column wash (20 mM imidazole); W2: second column wash (50 mM imidazole); W3: third column wash (75 mM imidazole); E1, E2, E3: elution with 500 mM imidazole; P: E2 fraction after desalting; P2: E1+E3+W4

The results obtained (Figure 3.11-13) are consistent with the previously obtained ones (Figures 3.8-10). With these results it was possible to define a successful and reproducible purification protocol for *Sa*SQR_WT. The resulting purified protein was used for the second round of crystallization screens, described later in this chapter (section 3.4).

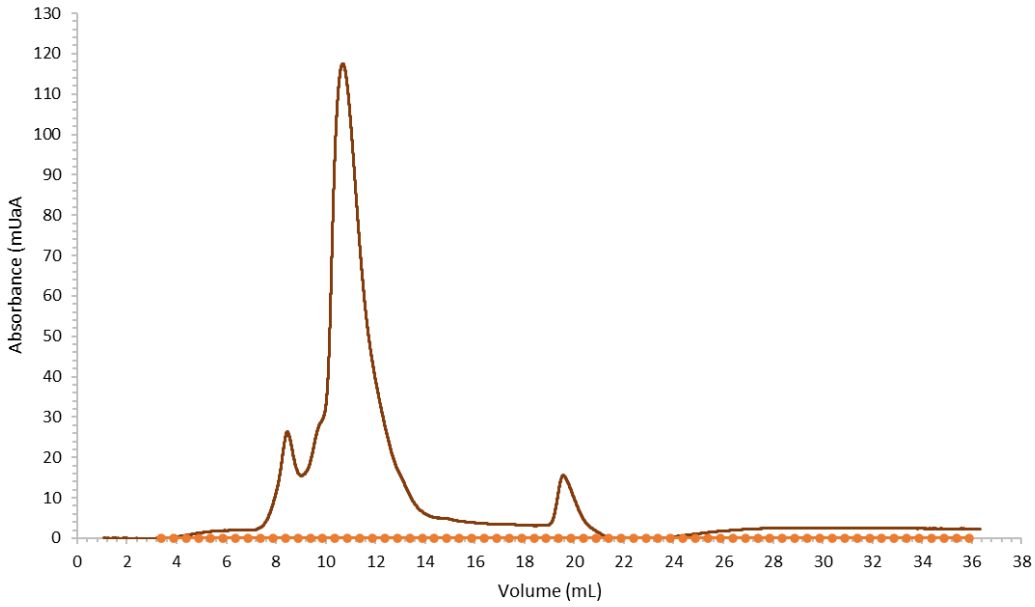

Figure 3.12: **Purification of *Sa*SQR by size exclusion chromatography.** Elution profile of *Sa*SQR using a Superdex 200 Increase 10/300 GL column on a ÄKTA Pure system. The protein of interest was injected in buffer S (Imidazole free) and elution was performed in the same buffer. Absorbance at 280 nm is represented in brown, and the fractions collected are represented with orange dots.

Later in this work, the same protocol was applied to the mutants *Sa*SQR_C133S and *Sa*SQR_C344S, but it did not yield such favorable results. The preliminary results for the purification of the mutants are reported in appendix A.5. Due to lack of time, the purification of the mutant proteins was not further optimized.

3.3. Thermal stability assays

Once a preliminary purification optimization was achieved, the next goal was to determine the ideal buffer conditions for *Sa*SQR purification and storage. For this, a series of thermostability assays were performed, comparing different commercial solutions.

A thermal shift assay (TSA) is a technique used to study the stability and interactions of proteins. It is a fast and efficient method that provides valuable information about the thermal denaturation of proteins and their binding to small molecules, ligands, or other proteins.[52]

In a TSA, the incubation of a fluorescent dye, such as SYPRO Orange (an environmentally sensitive fluorescent dye), is added to the protein sample. Increasing the temperature promotes protein unfolding with exposure of the protein's hydrophobic regions to which the dye binds.

The protein-dye mixture is subjected to a temperature gradient, typically by using a real-time PCR machine or a thermal cycler, and fluorescence is monitored over time.[53] Fluorescence increases as the protein undergoes thermal denaturation, as the core hydrophobic parts of the protein become exposed and more dye binds. The inflection point of the melting curve, known as the melting temperature (T_m), indicates the temperature at which half of the protein population is denatured.

In this case, a two-step approach was applied: first, the pH and chemical composition of the buffer were assessed, using the Durham pH screen; secondly, additives were explored using the RUBIC Additive screen, that contains small molecules that can affect protein folding, aggregation and stability (tables in appendix A.3).

In our own experience, it is commonly observed that, for successful crystallization, the melting temperature (T_m) of a protein must be in the range of 60 – 70 ºC. It is also suggested that the temperature shift for a good stability condition is around 20 ºC from room temperature to the T_m value.

Using the Durham pH screen, important information was collected about *Sa*SQR and its stability in particular buffers and chemical environments, allowing a more rational design of the next crystallization trials. Of all the conditions tested, none proved ideal in terms of the T_m and temperature shift, when compared to the buffer already in use. The results for the best 4 buffers are presented in figure 3.14.

3.3.1. Durham pH screen

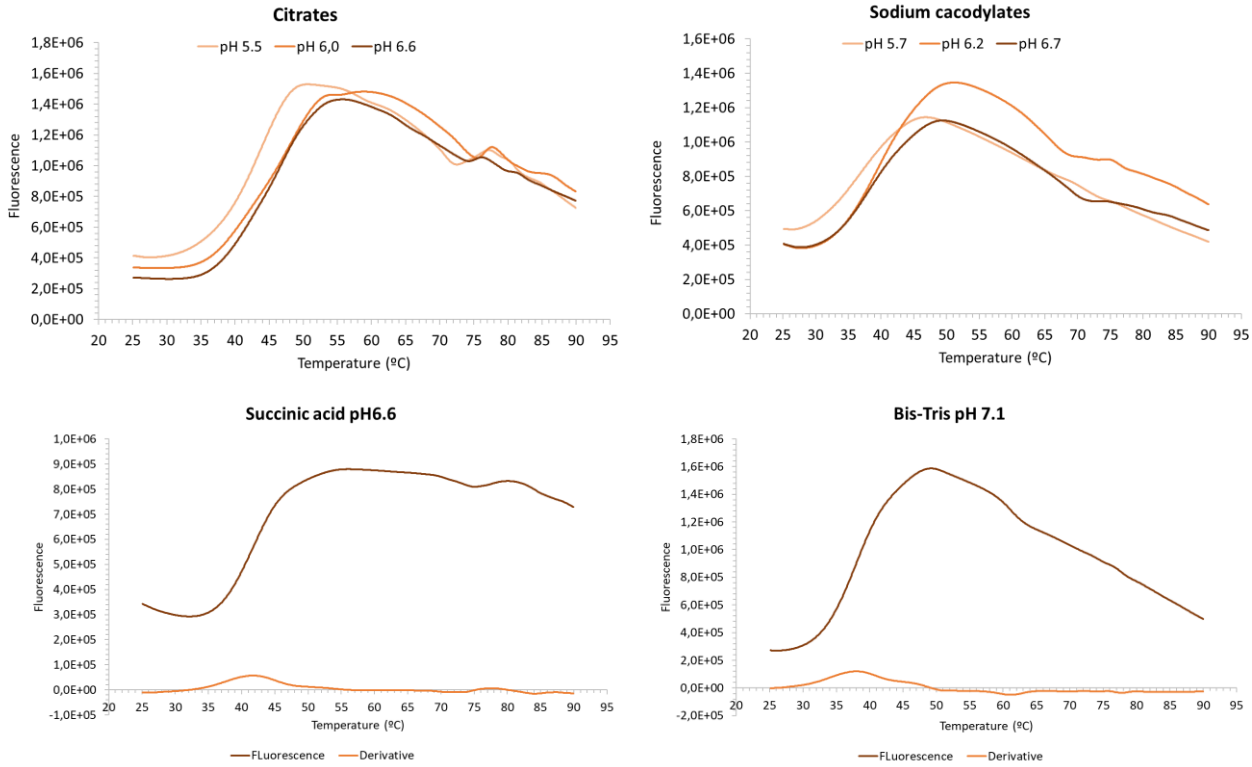

Figure 3.13: **Effect of buffer pH on *Sa*SQR thermal stability.** The Durham pH screen (Molecular Dimensions) was used In TSA assays to improve the purification and storage conditions of *Sa*SQR. The unfolding of 5 µg *Sa*SQR was monitored using 2x SYPRO Orange (Sigma Aldrich), in the range 20-95°C. The denaturation curves and respective melting temperatures for some selected conditions (indicated in

Table 3-1). Here are presented the best results obtained. In brown is the fluorescence curve obtained and in orange Its first derivative, which peak is the T_m value for each case.

Table 3-1: Values of T_m from the results obtained in the effect of buffer pH on *Sa*SQR thermal stability, via TSA.

Condition	Tris pH=8.0 (Control)	Citrate			Sodium cacodylate			Succinic acid	Bis-Tris
		pH 5.5	pH 6.0	pH 6.6	pH 5.7	pH 6.2	pH 6.7	pH 6.6	pH 7.1
T_m (ºC)	30.4	43.8	46.7	47.9	37.0	38.1	39.6	41.2	38.0

In both citrate buffers and sodium cacodylate buffers, the results demonstrate that there was a general trend of increasing T_m with increasing pH values and, therefore, increasing stability.

One of the highest T_m values obtained, 46.7 ºC, corresponds to the citrate buffer pH 6.6. However, the first derivative shows the presence of a double transition. This indicates that different parts of

the protein unfold at different moments instead of all at once. For this reason, this buffer was discarded. Although the denaturation curves for the sodium cacodylate buffers look good, the T_m values were lower than T_m for succinic acid at pH 6.6 (the highest value was around 39 ºC), hence these buffers were discarded as well.

The buffers Bis-Tris pH 7.1 and succinic acid pH 6.6 exhibited standard denaturation curves. The first yields a lower T_m value of 38 ºC, while the latter has a T_m value of 42 ºC. In another perspective, the succinic acid buffer has a higher initial fluorescence compared to Bis-Tris. This may indicate that the protein was already partially unfolded in the beginning, exposing core hydrophobic residues that bind the SYPRO-Orange dye. Taking all of this into consideration, we determined that the succinic acid buffer could be more suitable for *Sa*SQR stability, compared to the initial Tris pH 8.0 buffer being used.

In any case, as succinic acid does not present optimal conditions, especially in what concerns initial fluorescence, to further improve protein stability a second phase of tests was performed, using the RUBIC additive screen. The main goal for this experiment was to determine which small molecules could be added to the buffer to increase the T_m or reduce the initial fluorescence. The best results obtained are shown in figure 3.15. These encompass a mixture of two amino acids (arginine and glutamate), carbohydrate (glucose), a detergent (OG), a polyamine (the biotin) and the comparison between two reducing agents (DTT and TCEP).

3.3.2. RUBIC additive screen

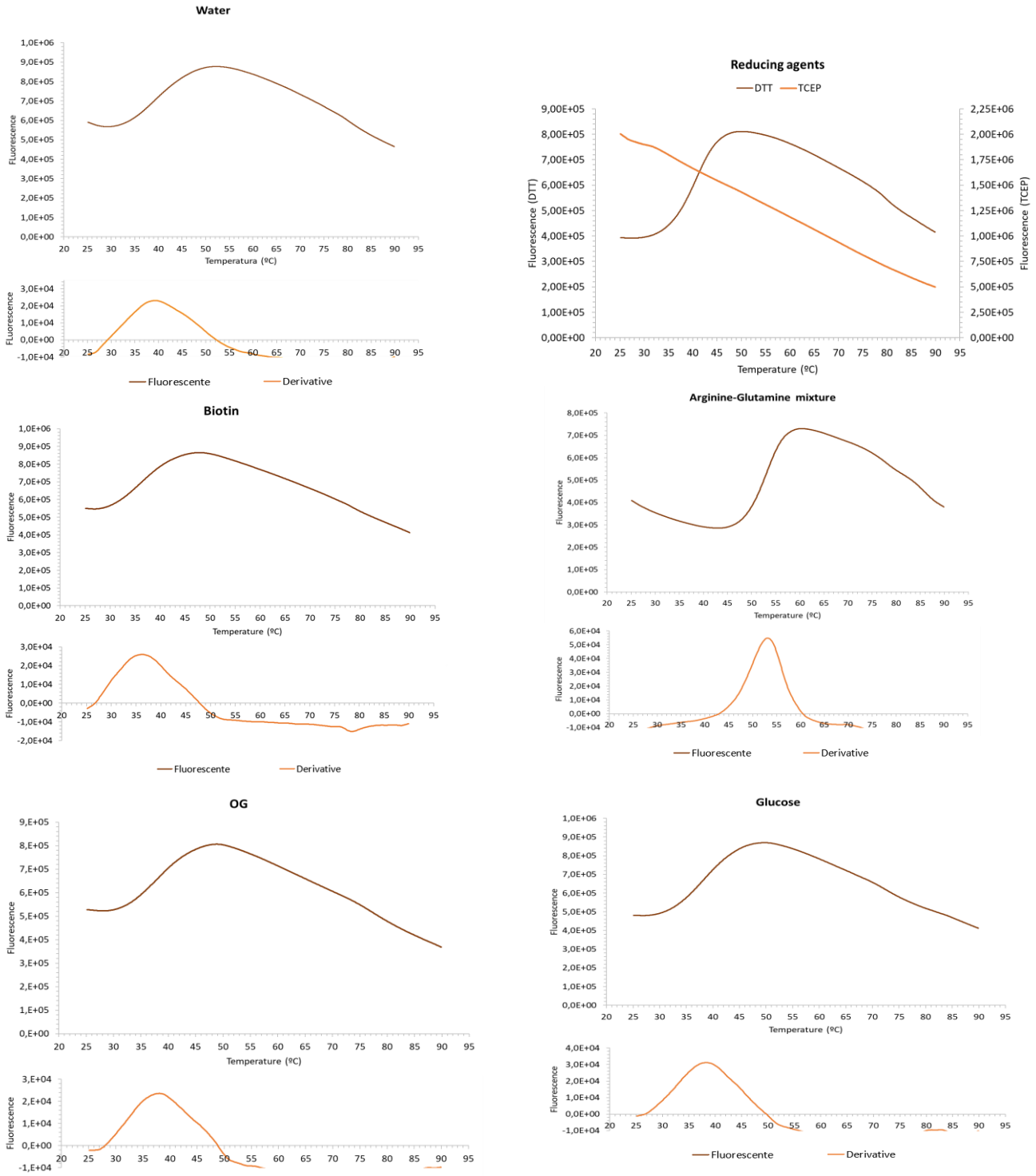

Figure 3.14: : **Effect of ionic strength and additives on *Sa*SQR thermal stability.** The RUBIC additive screen (Molecular Dimensions) was used to improve the purification and storage conditions for *Sa*SQR by TSA. Protein unfolding was monitored using 2x SYPRO Orange (Sigma Aldrich). The protein amount in all the assays was fixed to 5 µg. Data was collected at 1°C intervals from 20°C through 95°C. The denaturation curves and respective melting temperatures for some selected conditions (indicated in

Table 3-1Table 3-2). Here are presented the best results obtained. In brown is the fluorescence curve obtained and in orange Its first derivative, which peak is the T_m value for each case.

Table 3-2: Values of Tm obtained in TSA for the effect of Ionic strength and additives on SaSQR thermal stability.

Buffer	Tris pH 8.0						
Additive	Water (control)	5 mM TCEP	5 mM DTT	5 mM Biotin	500 mM Arginine-500 mM Glutamine mix.	1mM OG	25 mM Glucose
Tm (ºC)	39.4	-*	50.1	46.4	53.2**	37.9	38.2

Only the amino acid mixture proved to increase the T_m (up to 53.2 ºC), although maintaining the initial fluorescence high. All the other compounds reduced the initial fluorescence of the sample, but yielded low T_m values. Comparing the results from the amino acid mixture to those of each amino acid alone, it becomes clear that the T_m was only improved when both amino acids are present. Unfortunately, it was not possible to replicate this condition due to the low solubility of these amino acids.

Biotin, glucose, and the OG detergent showed similar T_m values (36.2, 38.2 and 38.0 ºC, respectively). Their presence in the buffer would help reduce the initial fluorescence in thermal stability assays, which is connected to a higher protein stability and could improve crystallization.

Following the suggestions of *Shen et al*., the reducing agent TCEP has been added to all buffers. However, from the TSA results it was possible to understand that TCEP is not suited for this protein. Nonetheless, another reducing agent, DTT, presents a "textbook-like" curve, with a T_m of 50.1 ºC. Therefore, TCEP was replaced by DTT in all buffers.

Based on these results, a new set of buffers was prepared, composed of succinic acid, DTT, glycerol, and FAD (buffer S2). However, when these buffers were used in the purification protocol previously defined, unexpected results were obtained. The Tm values for the S2 buffers were lower than the ones obtained before, with Tris buffer. For that reason, nano-DSF assays were performed with each buffer component separately and in different combinations, to understand their effect on protein stability.

Nano differential scanning fluorimetry (nanoDSF) is a technique used to analyze the stability and conformational changes of proteins. Like TSA, NanoDSF involves monitoring changes in protein fluorescence as a function of temperature.[54] However, contrarily to TSA, this technique relies on the intrinsic fluorescence of aromatic residues (such as tryptophane and tyrosine), which are commonly present in proteins.[55] Both techniques provide valuable information about the thermal stability of a protein, including its T_m. Due to its high cost, NanoDSF is mostly used to assess specific conditions that are considered the most promising in TSA.

During a nanoDSF experiment, a protein sample is subjected to a temperature ramp, from 20 to 95 ºC, while monitoring the fluorescence intensity. As the temperature increases, the protein undergoes

conformational changes, such as unfolding or aggregation, which can be detected by changes in the fluorescence signal. [10,11]

The most relevant results are presented in figure 3.16.

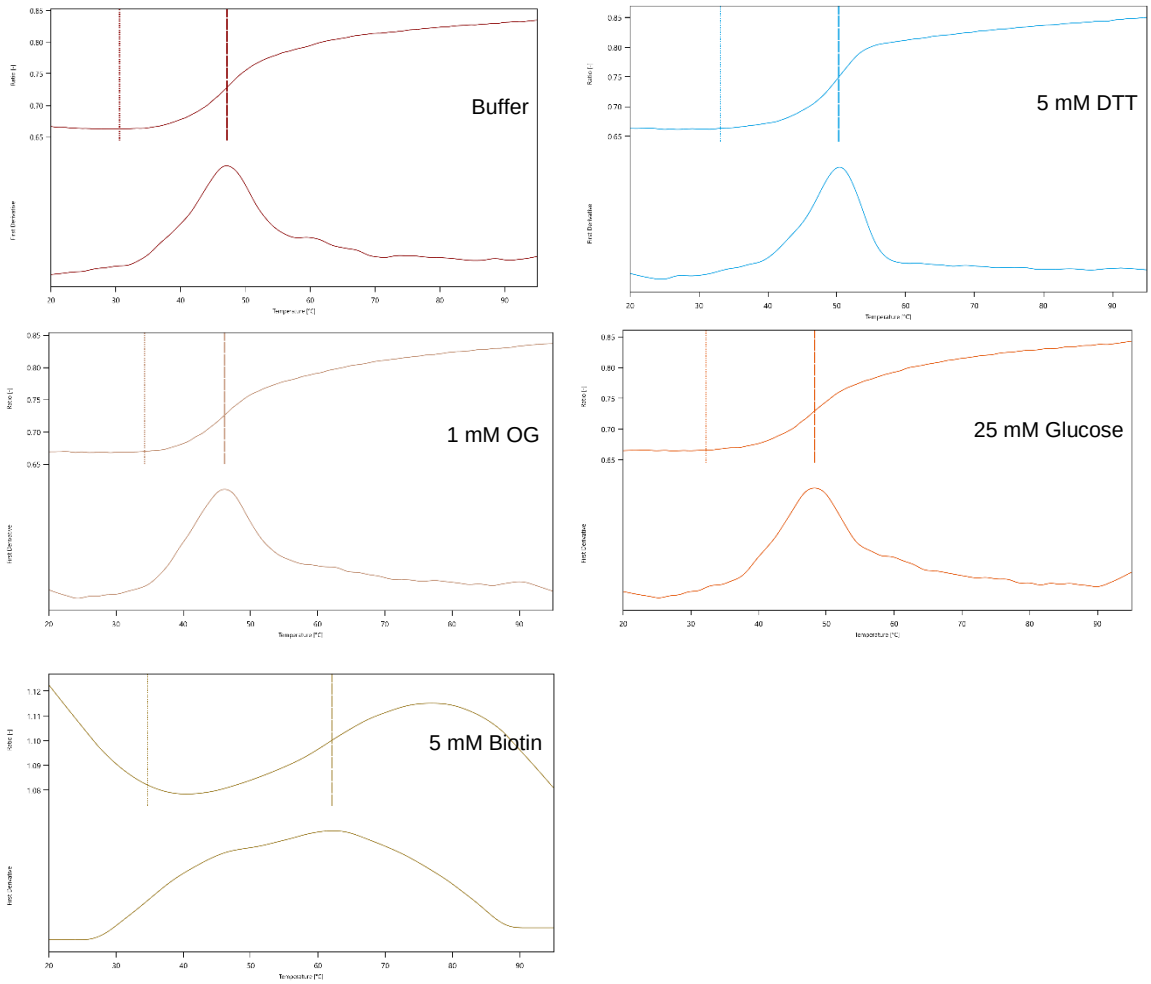

Figure 3.15: **NanoDSF analysis of S2 buffer components.** Thermal unfolding of 5 µg of *Sa*SQR was monitored at 330 nm in the range 20-95 °C. Each graph shows the Boltzmann (top) and first derivative (bottom) of denaturation curves. The denaturation curves and respective melting temperatures for some selected conditions are indicated in Table 3-3.

Table 3-3: Values of T_m obtained for the NanoDSF analysis of S2 buffer components.

Buffer	Tris pH 8.0				
Additive	- (control)	5 mM DTT	1mM OG	25 mM Glucose	5 mM Biotin
T_m (ºC)	47.1	50.3	46.2	48.3	62.1*

From these results it was possible to understand which compounds were favorable and which could be excluded. DTT, alone or combined with any other buffer compound, increases the T_m value. This further supports its inclusion in all buffers (replacing TCEP). Given its impact in protein stability, its concentration was increased from 1 to 5 mM.

The presence of glucose resulted in slow transitions, *i.e.,* the transition between the folded and unfolded state of the protein occurs during a larger temperature range, rendering the T_m values unreliable. Thus, glucose was removed from the buffers. The detergent OG was also removed from the buffers as its presence lead to lower T_m values, in all conditions. Lastly, the presence of biotin yielded not only a double and slow transition, but also a very high initial fluorescence. Although the later was in conflict with the previous TSA results, nanoDSF is more sensitive and therefore its results are more reliable. Thus, biotin was also removed.

New buffers were prepared based on these observations and another purification round was performed. However, the results for the new set of buffers presented worst data then the ones used before, since there was no separation of peaks in the chromatogram and the absorbance values were reduced, when in comparison with buffers used initially.

At this point, a step back was taken, and the succinic acid buffer was replaced by the initial Tris buffer, and new nanoDSF assays were performed. The new succinic acid buffer, the old Tris buffer, all the compounds of the Tris buffer separately, and the favorable RUBIC additives in the Tris buffer, separately and in different combinations, were tested. The most relevant results are presented in Figure 3.16 and Figure 3.17.

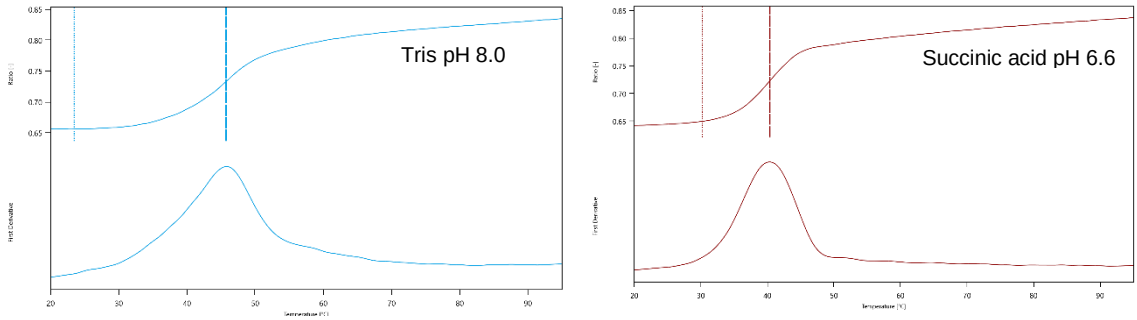

Figure 3.16: **NanoDSF analysis of *Sa*SQR in two different buffers.** Thermal unfolding of *Sa*SQR was monitored at 330 nm wavelength nm in the range 20-95 °C. The protein concentration in all the assays was fixed to 5 µg. Each graph shows the Boltzmann (top) and first derivative (bottom) of denaturation curves. The denaturation curves and respective melting temperatures for some selected conditions are indicated in Table 3-4.

Table 3-4: Values of T_m for nanoDSF analysis of two different buffers.

Condition	Tris pH 8.0	Succinic acid pH 6.6
T_m (ºC)	45.7	40.3

*Sa*SQR is more stable in the Tris buffer, with a T_m of 45.7 ºC, than in the succinic acid buffer, with a T_m of 40.3 ºC. Since the protein has a similar T_m in the Tris buffer and water, the addition of other compounds to the buffer may improve the T_m value to higher values, meaning better stability conditions for the protein.

To understand the role of each compound of the buffer in improving protein stability, the compounds of the buffer were tested separately and compared to the results obtained for Tris pH 8.0 (Figure 3.17).

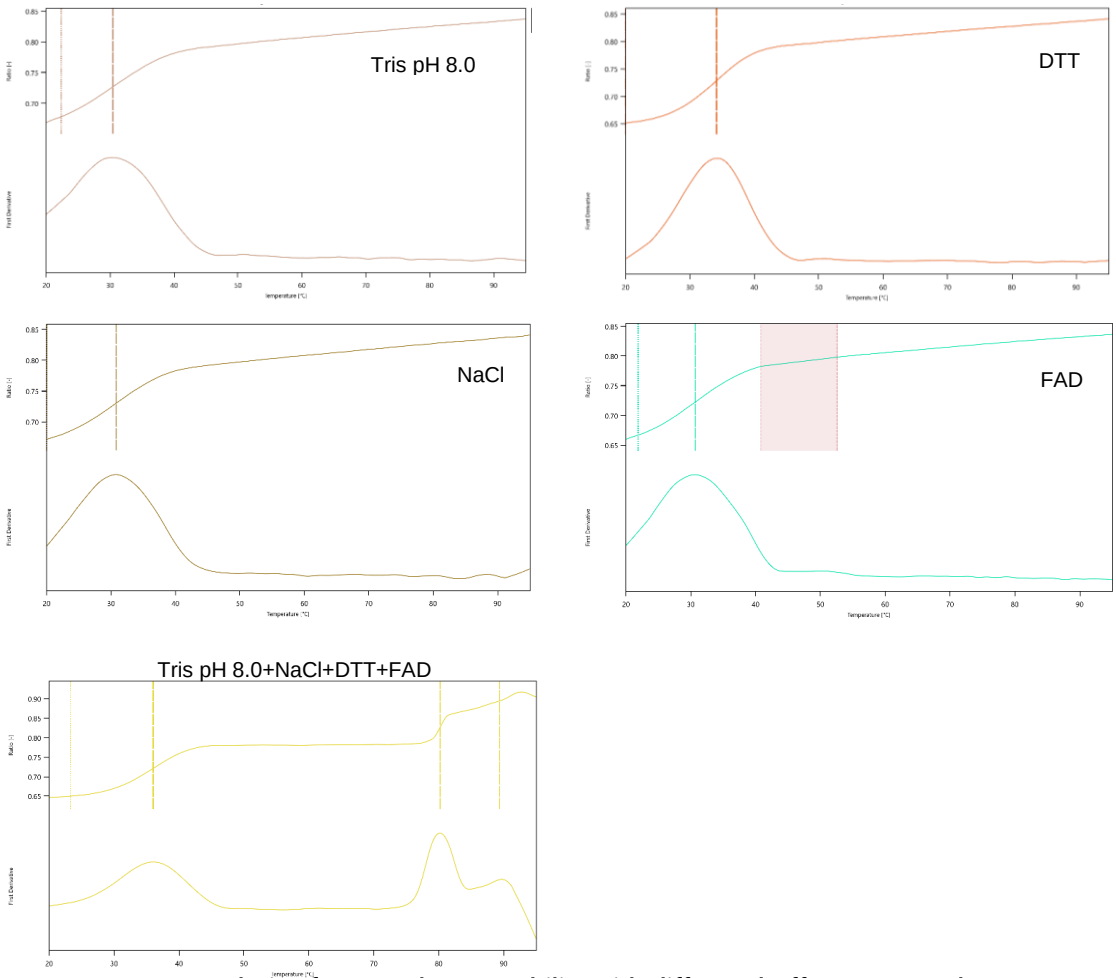

Figure 3.17: **NanoDSF analysis of *Sa*SQR thermostability with different buffer compounds.** Thermal unfolding of 5 µg of *Sa*SQR was monitored at 330 nm in the range 20-95 °C. Each graph shows the Boltzmann (up) and first derivative (low) melting profiles. The results presented are regarding to the TRIS HCl pH 8, DTT (top two panels), NaCl, FAD (middle panels), tested condition (Tris pH 8+NaCL+DTT+FAD) (bottom panel).

Table 3-5: Values of T_m obtained for the NanoDSF analysis of different buffer components.

Buffer	Tris pH 8.0				
Additive	None	5 mM DTT	150 mM NaCl	mM FAD	NaCl+DTT+FAD
T_m (ºC)	30.4	34.1	30.7	30.6	36.0

Addition of NaCl or FAD leads to denaturation curves similar to that of Tris, with T_m values around 30.6 ºC, meaning they have no effect. Addition of DTT increases the T_m value to 34.1 ºC, which further increases to 36.3 ºC when combined with FAD. This shows that FAD has an important role in the stability of the protein, but this was only visible when DTT was also present. A possible explanation is that DTT acts as a proton donor for FAD, stabilizing it, which in turn improves the protein stability. As for the small molecules that looked promising in the first nanoDSF trial, none proved favorable in this (Appendix A.6). The presence of the detergent OG led to T_m decrease up to 1.5 ºC relative to the best condition found (Tris, DTT, FAD, and NaCl). Glucose has no effect on the curve, hence, no effect on protein stability. Lastly, biotin shows a high initial fluorescence that compromises the reliability of the T_m values obtained.

Therefore, the protein buffer was changed to Tris pH 8.0 + DTT, NaCl and FAD. Since NaCl is required in all purification steps, and its presence does not affect *Sa*SQR thermostability, it was kept in the buffer formulation.

3.4. Crystallization trials

The main goal of this project was to obtain X-ray diffraction data to solve the structure of *Sa*SQR by X-ray Crystallography. For this, high amounts of highly pure protein are needed for crystallization trials. This was finally achieved after numerous optimizations.

Due to the lack of sufficient protein amounts, during the course of this project, 3 crystallization attempts were undertaken: 1 before testing the thermostability of the protein, and 2 after gaining knowledge of its thermostability.

For the first trial, two commercially available screens were initially tested: PACT Premier, a screen that tests the effect of pH, anions, and cations, using PEG as the precipitant; and BCS, a screen that covers a broader range of chemical space, while reducing the number of PEG variables, maintaining a large coverage of PEG space. These screens were chosen to ensure that the screening conditions covered a large part of the chemical space to enhance crystal formation. On average, crystallization screens were set up with 10 mg/mL of purified *Sa*SQR. For this first trial, no crystals were obtained from these screens, and most of the conditions resulted in heavy precipitates.

Due to the lack of favorable results in the first crystallization trial, and to improve the chances of crystallization, a Pre-Crystallization Trial (PCT) was done to determine the best range of protein concentration to use. It was determined that a concentration of 3 mg/mL was already high and caused heavy precipitation. By indication of the protocol, the concentration of protein was reduced to half.

In the second crystallization trial, which already had input from the thermostability data, 2 temperatures were tested, 4 and 20 ºC, and an additional commercial kit was used: JSCG+, a screen designed to maximize the coverage of the crystallization parameter space (buffers, salts, precipitants, and pH). The crystallization plates were monitored frequently during 6 weeks after their preparation. This time, after 19 days, a crystal was formed in 25% (w/v) PEG 1500 and 0.1 M SPG pH 6.0 (condition XXX of YYY screening), at 20 ºC.

Although SaSQR should look yellow under the stereomicroscope, the size of the crystal did not allow for a clear discernment between yellow and translucent. Since it is not possible to differentiate a salt crystal from a protein crystal by naked eye, this crystal was inspected under UV light in a SpectroLight 600 equipment. Due to the intrinsic fluorescence of amino acid residues like tryptophan, protein crystals fluoresce under UV light (Figure 3.18), allowing to discern between protein and salt crystals.

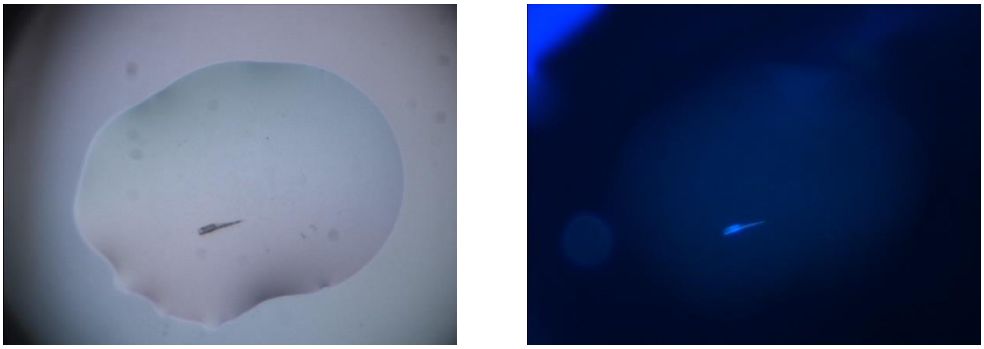

Figure 3.18: ***Sa*SQR protein crystal.** Crystallization condition A3 of the PACT Premier screen (25% w/v PEG 1500 and 0.1 M SPG pH 6.0). Drop volume of 200 nL with a 1:1 protein to precipitant ratio. The protein concentration used was 1.5 mg/ml in the buffer S. Left: crystallization drop illuminated with white light. Right: crystallization drop illuminated with UV light in the SpectroLight 600 (*Xtal concepts)*.

The crystallization condition in which the crystal appeared is composed by a succinic acid buffer, which correlates with the thermostability results obtained. The crystal obtained was stored in proper conditions and it is now waiting for the next synchrotron beam availability.

Since the crystallization plates set at 4 ºC showed no crystals after 6 weeks, they were moved to 20 ºC. However, they remained without the formation of protein crystals.

The third and last crystallization attempt was performed using three screens: PACT Premier screen, as used before; CSalt, a sparse matrix screen that uses a broad selection of salt-only solutions of varying concentrations and pH; the Morpheus screen, which was created ideally for membrane proteins, was also tested and it includes a range of pH, PEGs, salt additives and low molecular weight ligands to promote initial crystal formation. No crystals were observed for any of these conditions.

4

CONCLUSION AND FUTURE WORK

Sulfide:quinone oxidoreductases (SQRs) are proteins responsible for the first step of hydrogen sulfide (H_2S) detoxification. Acquiring structural knowledge about these proteins may give insights into their functional mechanisms. This can potentially unravel ways to modulate H_2S-driven bacterial resistance and mitigate bacterial infections, including the highly prevalent *Staphylococcus aureus* infections.

This project was primarily designed to characterize biochemically, structurally, and functionally the SQR protein from *S. aureus* (*Sa*SQR), and its two mutants *Sa*SQR_C133S and *Sa*SQR_C344S, that affect the active center and FAD-binding. The major goal of this work was to obtain crystallographic data for the three proteins and solve their structures. To do so, a protocol from gene to structure was developed for each protein.

Starting with the constructs kindly given by Giedroc's lab, the most suitable expression conditions were determined. A purification protocol for *Sa*SQR_WT was also developed and optimized from scratch, having as a starting point no effective protocol. Thermostability assays were performed using various buffers, pH, additives, and other conditions. From these results, it was possible to determine better buffer formulations to improve protein stability. Once sufficient amounts of purified protein were achieved, crystallization trials were pursued and a single *Sa*SQR crystal was obtained: 1.5 mg/mL of protein in the 0.1 M SPG pH 6.0 + 25% (w/v) PEG 1500 condition of the PACT Premier screen.

Therefore, the main purpose of this project was fulfilled: to go from a gene to a protein crystal. In future work, the crystallization conditions for *Sa*SQR_WT should be optimized in order to improve crystals for X-ray diffraction and structure determination.

The protocols for the expression and purification of the *Sa*SQR mutants (SaSQR_C133S and *Sa*SQR_C344S) must also be optimized, alongside with thermostability assays and crystallization screens, as this was not achieved within the timeframe of this thesis.

Ultimately, comparison of structural data from the wild type and the two mutants will allow the design of compounds with pharmacologic potential, and a complete biophysical and biochemical characterization of *Sa*SQR can be performed.

RESEARCH

1. Brewing up Earth's earliest life. *MIT News | Massachusetts Institute of Technology* https://news.mit.edu/2018/earths-first-biological-molecules-0409 (2018).
2. Olson, K. R. & Straub, K. D. The Role of Hydrogen Sulfide in Evolution and the Evolution of Hydrogen Sulfide in Metabolism and Signaling. *Physiology* **31**, 60–72 (2016).
3. Walsh, B. J. C., Brito, J. A. & Giedroc, D. P. Hydrogen Sulfide Signaling and Enzymology. in *Comprehensive Natural Products III* 430–473 (Elsevier, 2020). doi:10.1016/B978-0-12-409547-2.14699-2.
4. Filipovic, M. R., Zivanovic, J., Alvarez, B. & Banerjee, R. Chemical Biology of H_2S Signaling through Persulfidation. *Chem. Rev.* **118**, 1253–1337 (2018).
5. Neubeck, A. & Freund, F. Sulfur Chemistry May Have Paved the Way for Evolution of Antioxidants. *Astrobiology* **20**, 670–675 (2020).
6. SCHEELE, CARL WILHELM (1742 - 1786). http://rbx-exhibit2000.scs.illinois.edu/scheele.htm.
7. Hydrogen sulphide production in eggs and egg products as a result of heating - Germs - 1973 - Journal of the Science of Food and Agriculture - Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.2740240103.
8. International Union of Pure and Applied Chemistry. *IUPAC | International Union of Pure and Applied Chemistry* https://iupac.org/ (2023).
9. PubChem. Hydrogen Sulfide. https://pubchem.ncbi.nlm.nih.gov/compound/402.
10. Rickard, D. & Luther, G. W. Chemistry of Iron Sulfides. *Chem. Rev.* **107**, 514–562 (2007).
11. Madigan, M. T. *Brock biology of microorganisms*. (Pearson, 2015).
12. Gadalla, M. M. & Snyder, S. H. Hydrogen sulfide as a gasotransmitter. *J. Neurochem.* **113**, 14–26 (2010).

13. Warenycia, M. W. *et al.* Acute hydrogen sulfide poisoning. Demonstration of selective uptake of sulfide by the brainstem by measurement of brain sulfide levels. *Biochem. Pharmacol.* **38**, 973–981 (1989).

14. Kabil, O. & Banerjee, R. Characterization of Patient Mutations in Human Persulfide Dioxygenase (ETHE1) Involved in H2S Catabolism. *J. Biol. Chem.* **287**, 44561–44567 (2012).

15. Kimura, H. Hydrogen Sulfide as a Neuromodulator. *Mol. Neurobiol.* **26**, 013–020 (2002).

16. Cirino, G., Szabo, C. & Papapetropoulos, A. Physiological roles of hydrogen sulfide in mammalian cells, tissues, and organs. *Physiol. Rev.* (2022) doi:10.1152/physrev.00028.2021.

17. Lv, B. *et al.* Hydrogen sulfide and vascular regulation – An update. *J. Adv. Res.* **27**, 85–97 (2020).

18. Taylor, T. A. & Unakal, C. G. Staphylococcus Aureus. in *StatPearls* (StatPearls Publishing, 2023).

19. Lowy, F. D. Staphylococcus aureus Infections. *N. Engl. J. Med.* (1998).

20. Rasigade, J.-P. & Vandenesch, F. Staphylococcus aureus: A pathogen with still unresolved issues. *Infect. Genet. Evol.* **21**, 510–514 (2014).

21. Kluytmans, J., Belkum, A. V. & Verbrugh, H. Nasal Carriage of Staphylococcus aureus: Epidemiology, Underlying Mechanisms, and Associated Risks. *CLIN MICROBIOL REV* **10**, (1997).

22. Boucher, H. W. & Corey, G. R. Epidemiology of Methicillin-Resistant *Staphylococcus aureus*. *Clin. Infect. Dis.* **46**, S344–S349 (2008).

23. Wertheim, H. F. *et al.* Risk and outcome of nosocomial Staphylococcus aureus bacteraemia in nasal carriers versus non-carriers. *The Lancet* **364**, 703–705 (2004).

24. Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L. & Fowler, V. G. Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clin. Microbiol. Rev.* **28**, 603–661 (2015).

25. Lubna *et al.* Antimicrobial Usage and Detection of Multidrug-Resistant Staphylococcus aureus: Methicillin- and Tetracycline-Resistant Strains in Raw Milk of Lactating Dairy Cattle. *Antibiotics* **12**, 673 (2023).

26. Raygada, J. L. & Levine, D. P. Methicillin-Resistant Staphylococcus aureus: A Growing Risk in the Hospital and in the Community. *Am. Health* **2**, (2009).

27. Shen, J., Peng, H., Zhang, Y., Trinidad, J. C. & Giedroc, D. P. *Staphylococcus aureus sqr* Encodes a Type II Sulfide:Quinone Oxidoreductase and Impacts Reactive Sulfur Speciation in Cells. *Biochemistry* **55**, 6524–6534 (2016).

28. Grossoehme, N. *et al.* Control of Copper Resistance and Inorganic Sulfur Metabolism by Paralogous Regulators in Staphylococcus aureus. *J. Biol. Chem.* **286**, 13522–13531 (2011).

29. Luebke, J. L. *et al.* The CsoR -like sulfurtransferase repressor (CstR) is a persulfide sensor in *S taphylococcus aureus*. *Mol. Microbiol.* **94**, 1343–1360 (2014).

30. Shen, J. *et al. Staphylococcus aureus* CstB Is a Novel Multidomain Persulfide Dioxygenase-Sulfurtransferase Involved in Hydrogen Sulfide Detoxification. *Biochemistry* **54**, 4542–4554 (2015).

31. Higgins, K. A., Peng, H., Luebke, J. L., Chang, F.-M. J. & Giedroc, D. P. Conformational Analysis and Chemical Reactivity of the Multidomain Sulfurtransferase, *Staphylococcus aureus* CstA. *Biochemistry* **54**, 2385–2398 (2015).

32. Information on EC 1.8.5.4 - bacterial sulfide:quinone reductase - BRENDA Enzyme Database. https://www.brenda-enzymes.org/enzyme.php?ecno=1.8.5.4.

33. Jackson, M. R., Loll, P. J. & Jorns, M. S. X-Ray Structure of Human Sulfide:Quinone Oxidoreductase: Insights into the Mechanism of Mitochondrial Hydrogen Sulfide Oxidation. *Structure* **27**, 794-805.e4 (2019).

34. Sousa, F. M., Pereira, J. G., Marreiros, B. C. & Pereira, M. M. Taxonomic distribution, structure/function relationship and metabolic context of the two families of sulfide dehydrogenases: SQR and FCSD. *Biochim. Biophys. Acta BBA - Bioenerg.* **1859**, 742–753 (2018).

35. Knaff, D. B. & Buchanan, B. B. Cytochrome b and photosynthetic sulfur bacteria. *Biochim. Biophys. Acta BBA - Bioenerg.* **376**, 549–560 (1975).

36. Shahak, Y., Arieli, B., Binder, B. & Padan, E. Sulfide-dependent photosynthetic electron flow coupled to proton translocation in thylakoids of the cyanobacterium Oscillatoria limnetica. *Arch. Biochem. Biophys.* **259**, 605–615 (1987).

37. Arieli, B., Shahak, Y., Taglicht, D., Hauska, G. & Padan, E. Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in Oscillatoria limnetica. *J. Biol. Chem.* **269**, 5705–5711 (1994).

38. Theissen, U. Single Eubacterial Origin of Eukaryotic Sulfide:Quinone Oxidoreductase, a Mitochondrial Enzyme Conserved from the Early Evolution of Eukaryotes During Anoxic and Sulfidic Times. *Mol. Biol. Evol.* **20**, 1564–1574 (2003).

39. Brito, J. A. *et al.* Structural and Functional Insights into Sulfide:Quinone Oxidoreductase '. *Biochemistry* **48**, 5613–5622 (2009).

40. Bank, R. P. D. RCSB PDB - 3H8L: The first X-ray structure of a sulfide:quinone oxidoreductase: insights into sulfide oxidation mechanism. https://www.rcsb.org/structure/3H8L.

41. Marcia, M., Ermler, U., Peng, G. & Michel, H. The structure of *Aquifex aeolicus* sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc. Natl. Acad. Sci.* **106**, 9625–9630 (2009).

42. Bank, R. P. D. RCSB PDB - 3HYV: 3-D X-Ray structure of the sulfide:quinone oxidoreductase from the hyperthermophilic bacterium Aquifex aeolicus. https://www.rcsb.org/structure/3HYV.

43. A new structure-based classification of sulfide:quinone oxidoreductases - Marcia - 2010 - Proteins: Structure, Function, and Bioinformatics - Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.22665.

44. Landry, A. P. *et al.* A Catalytic Trisulfide in Human Sulfide Quinone Oxidoreductase Catalyzes Coenzyme A Persulfide Synthesis and Inhibits Butyrate Oxidation. *Cell Chem. Biol.* **26**, 1515-1525.e4 (2019).

45. GIEDROC GROUP: Chemical and Physical Biology of Infectious Disease. https://giedroc.lab.indiana.edu/.

46. Membrane Protein Crystallography Laboratory. https://www.itqb.unl.pt/research/biological-chemistry/membrane-protein-crystallography.

47. SupremeRun & LightRun Sequencing Services. https://eurofinsgenomics.eu/en/custom-dna-sequencing/gatc-services/.

48. Sequence map of pHis-parallel1-NLSH2BCas9 plasmid. https://www.addgene.org/112065/ (2023).

49. increase protein solubility and expression in E. coli | Agilent. https://www.agilent.com/en/product/protein-expression/competent-cells-for-difficult-protein-expression/insoluble-proteins/arcticexpress-%28de3%29-competent-cells-232939.

50. Kroemer, T. How Does IPTG Induction Work? https://goldbio.com/articles/article/how-does-iptg-induction-work#:~:text=IPTG%20or%20Isopropyl%20%CE%B2%2DD,genes%20in%20the%20lac%20operon.

51. Expasy ProtParam tool. https://web.expasy.org/cgi-bin/protparam/protparam.

52. Thermal Shift Assay Services | Reaction Biology. https://www.reactionbiology.com/services/biophysical-assays/thermal-shift-assay-tsa.

53. Huynh, K. & Partch, C. L. Current Protocols in Protein Science. *Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al* **79**, 28.9.1-28.9.14 (2015).

54. Nano differential scanning fluorimetry (nanoDSF / nDSF). https://www.coriolis-pharma.com/analytical-services/higher-order-structure-analysis/nano-differential-scanning-fluorimetry-nanodsf.

55. Magnusson, A. O. *et al.* nanoDSF as screening tool for enzyme libraries and biotechnology development. *Febs J.* **286**, 184–204 (2019).

APPENDIX

A.1 Theoretical Information of *Sa*SQR obtained by ExPasy[51]

Number of amino acids: 421

Molecular weight: 47780.40

Theoretical pI: 8.60

Formula: $C_{2145}H_{3318}N_{578}O_{629}S_{16}$

A.2 Sequences

Sequence alignments were obtained using the T7 promotor. In red, there are highlited the residues that give rise to the mutation, changing the wild type protein sequence to a mutant protein sequence.

A.2.1 Original sequence of SaSQR_WT with 6xHis-tag

HHHHHH DYDI PTTENLYFQG A MEKMNKHYQ IVIIGGGTAG VTVASRLLRK NQNLKEKIAI IDPADHHYYQ PLWTLVGAGV SSLKSSRKDM ESVIPEGANW IKQAVSSFQP ENNSVILGDN TVVYYDFLVV APGLQINWSS IKGLKENIGK NGVCSNYSPD YVNETWNQIS NFKQGNAIFT HPNTPIKCGG APMKIMYLAE DYFRKHKIRS NANVIYATPK DALFDVGKYN KELERIVEER NITVNYNYNL VEIDGDKKVA TFEHIKAYDR KTISYDMLHV TPPMGPLDVV KESTLSDSEG WVDVNPTTLQ HKSYSNVFAL GDASNVPTSK TGAAIRKQAP IVANNLLQVM NNQMLTHHYD GYTSCPIVTG YNRLILAEFD YNKNTKETMP FNQAKERRSM YIFKKDLLPK MYWYGMLKGL

A.2.2 Sequencing alignment of *Sa*SQR_WT

```
SaSQR_WT_T7       MSYYHHHHHHDYDIPTTENLYFQGAMEKMNKHYQIVIIGGGTAGVTVASRLLRKNQNLKE
SaSQR_WT_T7term   ------------------------------------------------------------
SaSQR             ----HHHHHHDYDIPTTENLYFQGAMEKMNKHYQIVIIGGGTAGVTVASRLLRKNQNLKE

SaSQR_WT_T7       KIAIIDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVI
SaSQR_WT_T7term   ------------------------------------------------------------
SaSQR             KIAIIDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVI
```

```
SaSQR_WT_T7       LGDNTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN
SaSQR_WT_T7term   --------MIFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN
SaSQR             LGDNTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN


SaSQR_WT_T7       AIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI
SaSQR_WT_T7term   AIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI
SaSQR             AIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI


SaSQR_WT_T7       VEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS
SaSQR_WT_T7term   VEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS
SaSQR             VEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS


SaSQR_WT_T7       DSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT
SaSQR_WT_T7term   DSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT
SaSQR             DSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT


SaSQR_WT_T7       HHYDGYTSCPIGTGYNRLILAEFDYNKILRKHAVYQA-KERRV-----------------
SaSQR_WT_T7term   HHYDGYTSCPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGM
SaSQR             HHYDGYTSCPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGM


SaSQR_WT_T7          -----
SaSQR_WT_T7term      LKGLI
SaSQR                LKGLI
```

A.2.3 Sequence alignment of SaSQR_C133S

```
SaSQR_C133S_T     ------------------------MNKHYQIVIIGGGTAGVTVASRLLRKNQNLKEKIAI
SaSQR_C133S_T7term------------------------------------------------------------
SaSQR             HHHHHHDYDIPTTENLYFQGAMEKMNKHYQIVIIGGGTAGVTVASRLLRKNQNLKEKIAI


SaSQR_C133S_T7    IDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVILGDN
SaSQR_C133S_T7term----------------------------MESVIPEGANWIKQAVSSFQPENNSVILGDN
SaSQR             IDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVILGDN


SaSQR_C133S_T7    TVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVSSNYSPDYVNETWNQISNFKQGNAIFT
SaSQR_C133S_T7termTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVSSNYSPDYVNETWNQISNFKQGNAIFT
SaSQR             TVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGNAIFT


SaSQR_C133S_T7    HPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERIVEER
SaSQR_C133S_T7termHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERIVEER
SaSQR             HPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERIVEER


SaSQR_C133S_T7    NITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLSDSEG
SaSQR_C133S_T7termNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLSDSEG
SaSQR             NITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLSDSEG


SaSQR_C133S_T7    WVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLTHHYD
SaSQR_C133S_T7termWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLTHHYD
SaSQR             WVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLTHHYD


SaSQR_C133S_T7    GYTSCPIVTGYK-----------------------------------------------
SaSQR_C133S_T7termGYTSCPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGMLKGL
```

```
SaSQR             GYTSCPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGMLKGL


SaSQR_C133S_T7          -
SaSQR_C133S_T7term      I
SaSQR                   I
```

A.2.4 Sequence alignment of *Sa*SQR_C344S

```
SaSQR_C344S_T7    MSYYHHHHHHDYDIPTTENLYFQGAMEKMNKHYQIVIIGGGTAGVTVASRLLRKNQNLKE
SaSQR_C344S_T7term------------------------------------------------------------
SaSQR             ----HHHHHHDYDIPTTENLYFQGAMEKMNKHYQIVIIGGGTAGVTVASRLLRKNQNLKE


SaSQR_C344S_T7    KIAIIDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVI
SaSQR_C344S_T7term---------------------------------MESVIPEGANWIKQAVSSFQPENNSVI
SaSQR             KIAIIDPADHHYYQPLWTLVGAGVSSLKSSRKDMESVIPEGANWIKQAVSSFQPENNSVI


SaSQR_C344S_T7    LGDNTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN
SaSQR_C344S_T7termLGDNTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN
SaSQR             LGDNTVVYYDFLVVAPGLQINWSSIKGLKENIGKNGVCSNYSPDYVNETWNQISNFKQGN


SaSQR_C344S_T7    AIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI
SaSQR_C344S_T7termAIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI
SaSQR             AIFTHPNTPIKCGGAPMKIMYLAEDYFRKHKIRSNANVIYATPKDALFDVGKYNKELERI

SaSQR_C344S_T7    VEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS
SaSQR_C344S_T7termVEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS
SaSQR             VEERNITVNYNYNLVEIDGDKKVATFEHIKAYDRKTISYDMLHVTPPMGPLDVVKESTLS


SaSQR_C344S_T7    DSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT
SaSQR_C344S_T7termDSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT
SaSQR             DSEGWVDVNPTTLQHKSYSNVFALGDASNVPTSKTGAAIRKQAPIVANNLLQVMNNQMLT


SaSQR_C344S_T7    HHYDGYTSSPIVTGNNRLILAEFDYNKILKKQCRLIRPRT--------------------
SaSQR_C344S_T7termHHYDGYTSSPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGM
SaSQR             HHYDGYTSCPIVTGYNRLILAEFDYNKNTKETMPFNQAKERRSMYIFKKDLLPKMYWYGM


SaSQR_C344S_T7          -----
SaSQR_C344S_T7term      LKGLI
SaSQR                   LKGLI
```

A.3 RUBBIC Additive screen

	Concentration during assay		Concentration during assay
A1	100 % Ultrapure water	**E1**	1 mM CHAPS
A2	100 mM Sodium acetate trihydrate	**E2**	1 mM CHAPSO
A3	100 mM Calcium acetate hydrate	**E3**	1 mM OG
A4	100 mM Potassium acetate	**E4**	1 mM DM
A5	100 mM Ammonium acetate	**E5**	1 mM DDM
A6	100 mM Sodium sulfate	**E6**	25 mM Monosaccharides mix
A7	100 mM Magnesium sulfate	**E7**	25 mM D-Glucose
A8	100 mM Potassium sulfate	**E8**	25 mM Sucrose
A9	100 mM Ammonium sulfate	**E9**	25 mM Maltose
A10	100 mM Sodium phosphate monobasic	**E10**	50 mM Carboxylic acids mix
A11	100 mM Sodium phosphate dibasic	**E11**	50 mM L-Proline
A12	100 mM Potassium phosphate	**E12**	50 mM Glycine
B1	100 mM Potassium phosphate dibasic	**F1**	50 mM L-Glutamic acid monosodium salt
B2	100 mM Sodium tartrate dibasic	**F2**	500 mM L-Glutamic acid monosodium salt
B3	100 mM Sodium citrate tribasic	**F3**	50 mM L-Arginine
B4	100 mM Sodium malonate dibasic	**F4**	500 mM L-Arginine
B5	100 mM Sodium nitrate	**F5**	50 mM L-Glutamic acid /50 mM L-Arginine
B6	100 mM Sodium formate	**F6**	500 mM L-Glutamic acid /500 mM L-Arginine
B7	100 mM Potassium formate	**F7**	50 mM Gly-Gly-Gly
B8	100 mM Sodium fluoride	**F8**	5 mM Oxaloacetic acid
B9	100 mM Potassium fluoride	**F9**	5 % v/v Dimethyl sulfoxide
B10	100 mM Ammonium fluoride	**F10**	5 % v/v Ethylene glycol
B11	100 mM Lithium chloride	**F11**	5 % v/v Glycerol
B12	100 mM Sodium chloride	**F12**	20 % v/v Glycerol
C1	100 mM Potassium chloride	**G1**	5 % v/v PEG 400
C2	100 mM Ammonium chloride	**G2**	5 % w/v PEG 1000
C3	100 mM Sodium iodide	**G3**	5 % w/v PEG 3350
C4	100 mM Potassium iodide	**G4**	5 mM DTT
C5	100 mM Sodium bromide	**G5**	5 mM TCEP
C6	1 mM Magnesium chloride	**G6**	5 mM Biotin
C7	1 mM Calcium chloride	**G7**	5 mM Betaine hydrochloride
C8	1 mM Manganese(II) chloride	**G8**	5 mM Coenzyme A
C9	1 mM Nickel(II) chloride	**G9**	5 mM Nicotinic acid
C10	1 mM Iron(III) chloride	**G10**	1 mM Spermidine
C11	1 mM Zinc chloride	**G11**	1 mM Spermine tetrahydrochloride
C12	1 mM Cobalt(II) chloride hexahydrate	**G12**	1 mM Sarcosine
D1	5 mM EDTA	**H1**	20 µM Deoxyribonucleic acid
D2	5 mM EGTA	**H2**	1 mM ATP/ 1mM Magnesium chloride
D3	0.1 M Urea	**H3**	1 mM ATPγS/ 1mM Magnesium chloride
D4	0.5 M Urea	**H4**	1 mM cAMP/ 1mM Magnesium chloride
D5	1 M Urea	**H5**	1 mM GTP/ 1mM Magnesium chloride
D6	2 M Urea	**H6**	1 mM GTPγS/ 1mM Magnesium chloride
D7	4 M Urea	**H7**	1 mM cGMP/ 1mM Magnesium chloride
D8	150 mM Guanidine hydrochloride	**H8**	1 mM NADH/ 1mM Magnesium chloride
D9	500 mM Guanidine hydrochloride	**H9**	1 mM NADPH/ 1mM Magnesium chloride
D10	1 mM NDSB 195	**H10**	5 mM Polyethyleneimine 800
D11	1 mM NDSB 201	**H11**	200 mM Imidazole
D12	1 mM Fos-Choline-12	**H12**	400 mM Imidazole

A.4 Durham pH Screen (Molecular Dimensions)

	Concentration during assay		Concentration during assay
A1	100 % Ultrapure water	**E1**	100 mM Sodium phosphate pH 6.3
A2	100 % Ultrapure water	**E2**	100 mM Sodium phosphate pH 6.8
A3	4 M Urea	**E3**	100 mM Sodium phosphate pH 7.3
A4	100 mM Citrate pH 4.1	**E4**	100 mM PIPES pH 6.3
A5	100 mM Citrate pH 4.6	**E5**	100 mM PIPES pH 6.8
A6	100 mM Citrate pH 5.1	**E6**	100 mM PIPES pH 7.3
A7	100 mM Acetic acid pH 4.2	**E7**	100 mM Imidazole pH 6.6
A8	100 mM Acetic acid pH 4.7	**E8**	100 mM Imidazole pH 7.1
A9	100 mM Acetic acid pH 5.2	**E9**	100 mM Imidazole pH 7.6
A10	100 mM Succinic acid pH 4.4	**E10**	100 mM MOPS pH 6.6
A11	100 mM Succinic acid pH 4.9	**E11**	100 mM MOPS pH 7.1
A12	100 mM Succinic acid pH 5.4	**E12**	100 mM MOPS pH 7.6
B1	100 mM DL-Malic acid pH 4.3	**F1**	100 mM Bis-Tris propane pH 6.6
B2	100 mM DL-Malic acid pH 4.8	**F2**	100 mM Bis-Tris propane pH 7.1
B3	100 mM DL-Malic acid pH 5.3	**F3**	100 mM Bis-Tris propane pH 7.6
B4	100 mM L-Tartaric acid pH 4.3	**F4**	100 mM HEPES pH 7.0
B5	100 mM L-Tartaric acid pH 4.8	**F5**	100 mM HEPES pH 7.5
B6	100 mM L-Tartaric acid pH 5.3	**F6**	100 mM HEPES pH 8.0
B7	100 mM Propionic acid pH 4.3	**F7**	100 mM Tricine pH 7.5
B8	100 mM Propionic acid pH 4.8	**F8**	100 mM Tricine pH 8.0
B9	100 mM Propionic acid pH 5.3	**F9**	100 mM Tricine pH 8.5
B10	100 mM Malonic acid pH 5.2	**F10**	100 mM EPPS pH 7.5
B11	100 mM Malonic acid pH 5.7	**F11**	100 mM EPPS pH 8.0
B12	100 mM Malonic acid pH 6.2	**F12**	100 mM EPPS pH 8.5
C1	100 mM Citrate pH 5.5	**G1**	100 mM Tris pH 7.7
C2	100 mM Citrate pH 6.0	**G2**	100 mM Tris pH 8.2
C3	100 mM Citrate pH 6.5	**G3**	100 mM Tris pH 8.7
C4	100 mM Succinic acid pH 5.6	**G4**	100 mM BICINE pH 7.7
C5	100 mM Succinic acid pH 6.1	**G5**	100 mM BICINE pH 8.2
C6	100 mM Succinic acid pH 6.6	**G6**	100 mM BICINE pH 8.7
C7	100 mM MES pH 5.6	**G7**	100 mM TAPS pH 7.9
C8	100 mM MES pH 6.1	**G8**	100 mM TAPS pH 8.4
C9	100 mM MES pH 6.6	**G9**	100 mM TAPS pH 8.9
C10	100 mM Maleic acid pH 5.7	**G10**	100 mM Bis-Tris propane pH
C11	100 mM Maleic acid pH 6.2	**G11**	100 mM Bis-Tris propane pH 9.0
C12	100 mM Maleic acid pH 6.7	**G12**	100 mM Bis-Tris propane pH 9.5
D1	100 mM Sodium cacodylate pH 5.7	**H1**	100 mM Boric acid pH 8.6
D2	100 mM Sodium cacodylate pH 6.2	**H2**	100 mM Boric acid pH 9.1
D3	100 mM Sodium cacodylate pH 6.7	**H3**	100 mM Boric acid pH 9.6
D4	100 mM ADA pH 6.1	**H4**	100 mM CHES pH 8.8
D5	100 mM ADA pH 6.6	**H5**	100 mM CHES pH 9.3
D6	100 mM ADA pH 7.1	**H6**	100 mM CHES pH 9.8
D7	100 mM Bis-Tris pH 6.1	**H7**	100 mM Glycine pH 9.2
D8	100 mM Bis-Tris pH 6.6	**H8**	100 mM Glycine pH 9.7
D9	100 mM Bis-Tris pH 7.1	**H9**	100 mM Glycine pH 10.2
D10	100 mM ACES pH 6.3	**H10**	100 mM CAPS pH 9.9
D11	100 mM ACES pH 6.8	**H11**	100 mM CAPS pH 10.4
D12	100 mM ACES pH 7.3	**H12**	100 mM CAPS pH 10.9

A.5 Preliminary results for mutant purifications

A.5.1 *Sa*SQR_C133S

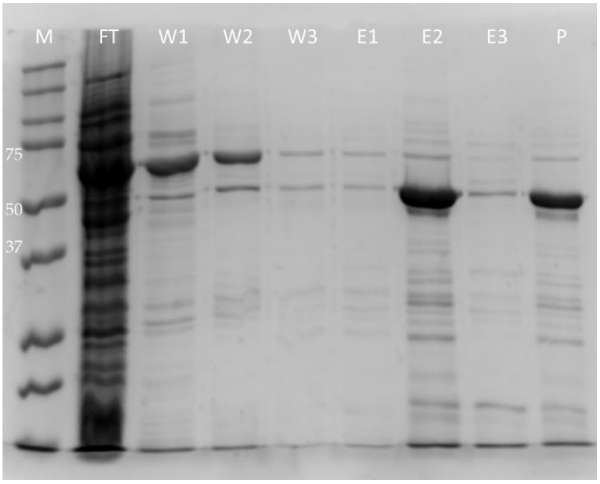

Figure A 1: **SDS-PAGE analysis of the efficiency of the affinity chromatography purification for SaSQR_C133S mutant**. *Sa*SQR_C133S was purified using a gravity-flow HisGraviTrap column. Samples were collected during elution and were analyzed using a 12% polyacrylamide run at 180 V: precision plus protein dual color standards (Biorad); FT: sample application flow-through; W1: first column wash (20 mM imidazole); W2: second column wash (50 mM imidazole); W3: third column wash (75 mM imidazole); E1, E2, E3: elution with 500 mM imidazole; P: E2 fraction after desalting.

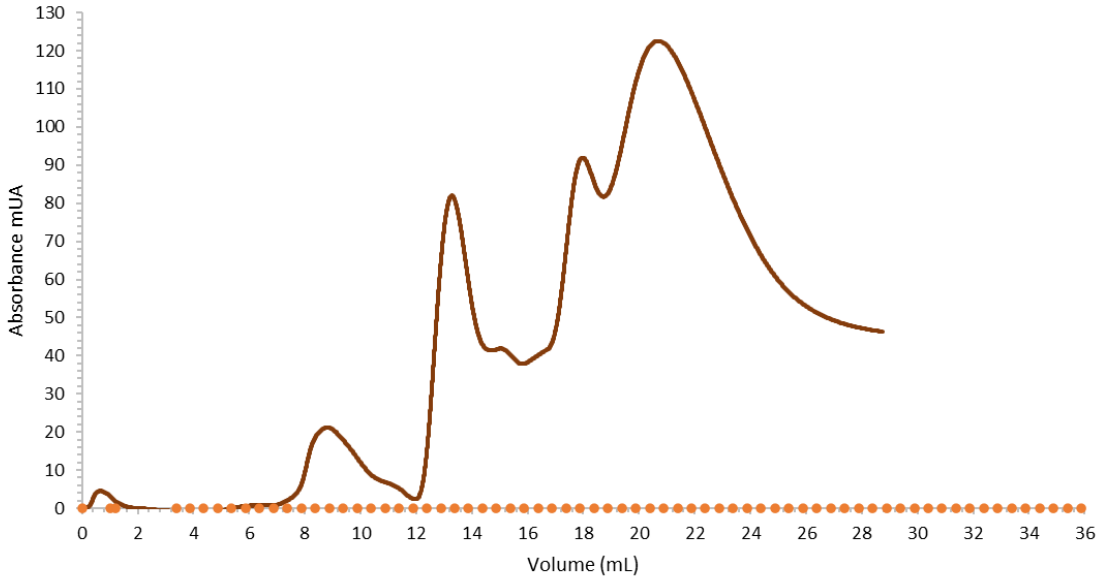

Figure A 2: **Purification of *Sa*SQR_C133S by size exclusion chromatography.** Elution profile of *Sa*SQR_C133S using a Superdex 200 Increase 10/300 GL column on a ÄKTA Pure system. The protein of interest was injected in buffer S (Imidazole free) and elution was performed in the same buffer. Absorbance at 280 nm is represented in brown, and the fractions collected are represented with orange dots.

A.5.2 *Sa*SQR_C344S

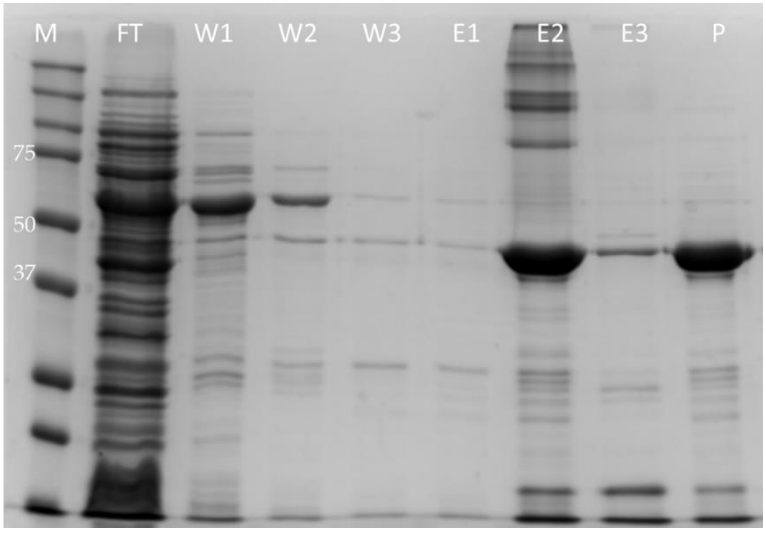

Figure A 3: **SDS-PAGE analysis of the efficiency of the affinity chromatography purification for SaSQR_C344S mutant**. *Sa*SQR_C344S was purified using a gravity-flow HisGraviTrap column. Samples were collected during elution and were analyzed using a 12% polyacrylamide run at 180 V. M: precision plus protein dual color standards (Biorad); FT: sample application flow-through; W1: first column wash (20 mM imidazole); W2: second column wash (50 mM imidazole); W3: third column wash (75 mM imidazole); E1, E2, E3: elution with 500 mM imidazole; P: E2 fraction after desalting.

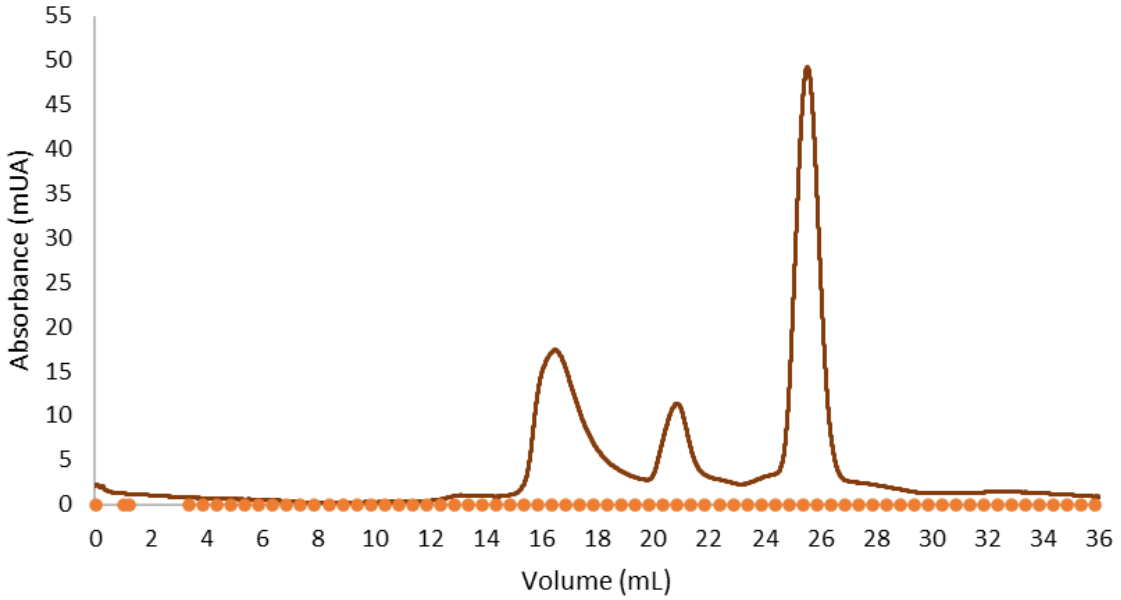

Figure A 4: **Purification of *Sa*SQR_C344S by size exclusion chromatography.** Elution profile of *Sa*SQR_C344S using a Superdex 200 Increase 10/300 GL column on a ÄKTA Pure system. The protein of interest was injected in buffer S (Imidazole free) and elution was performed in the same buffer. Absorbance at 280 nm is represented in brown, and the fractions collected are represented with orange dots.

A.6 Tabel of all T_m values

Conditions	Tm (ºC)
Durham pH screen	
Citrates	43.8 < 46.7 < 47.9
Sodium Cacodylates	37.0 < 38.1 < 39.6
Succinic acid pH 6.6	41.2
Bis-Tris pH 7.1	38
RUBIC additive screen	
Water	39.4
Reducing agents	**TCEP: - ** DTT: 50.1**
Biotine	46.4
Arginine-Glutamine	53.2*
OG	37.9
Glucose	38.2
**No Tm value was obtained.	
NanoDSF of S2 buffer components	
Buffer	47.1
DTT	50.3
OG	46.2
Glucose	48.3
Biotin	62.1*
*Slow transition and high initial fluorescence lead to unreliable results	
NanoDSF analysis of *Sa*SQR in two different buffers	
Tris pH 8.0	45.7
Succinic acid pH 6.6	40.3
NanoDSF analysis of *Sa*SQR thermostability with different buffer compounds	
TRIS HCl pH 8.0	30.4
DTT	34.1
NaCl	30.7
FAD	30.6
Tris pH 8.0 + NaCl + DTT + FAD	36.0

Structural and functional studies on enzymes involved in microbial hydrogen sulfide metabolism

Rafaela Almeida Cordeiro

UNIVERSIDADE NOVA
DE LISBOA

2023