

Structural and functional studies of enzymes

Tatiana Vieira Van der Kellen

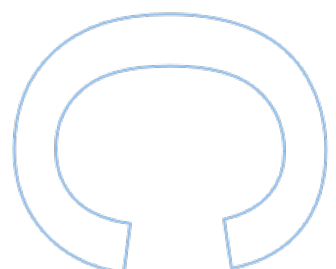
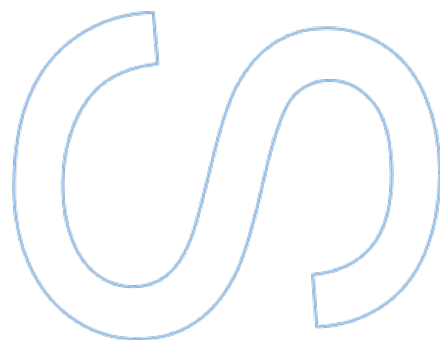
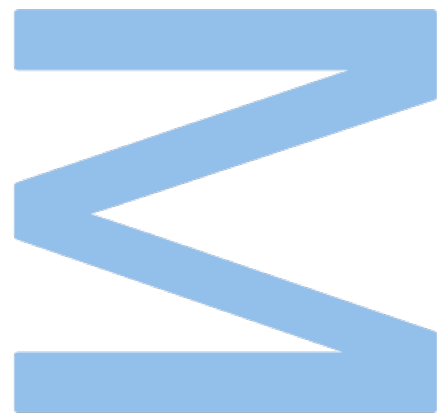
Masters in Biochemistry

Department of Chemistry and Biochemistry

Faculty of Sciences of University of Porto and Institute for Health

Research and Innovation, i3S

2024



Structural and functional studies of enzymes

Tatiana Vieira Van der Kellen

Masters in Biochemistry

Department of Chemistry and Biochemistry

Faculty of Sciences of University of Porto and Institute for Health

Research and Innovation, i3S

2024

Supervisor

Joana A. Santos, PhD, i3S

Co-Supervisor

Pedro Pereira, PhD, i3S

Co-supervisor

Prof. Dr. Luís Gales, Associate Professor, ICBAS



Acknowledgements

Primeiro, gostaria de agradecer ao Dr. Pedro Pereira por me ter recebido na sua equipa de investigação “Macromolecular Structure” e por todas as oportunidades que me ofereceu ao longo deste projeto. Foi sem dúvida um ano muito enriquecedor, onde aprendi muito.

Um agradecimento com muito carinho à minha orientadora Joana Santos por me ter acompanhado e guiado neste trabalho. Queria agradecer por todos os ensinamentos, conselhos e valores que me deu ao longo deste ano, não só a nível científico, como também a nível pessoal, e que certamente levarei para minha vida futura. Obrigada por nunca teres desistido de mim, especialmente nos ensaios enzimáticos que tanto me deram cabo da cabeça, e por teres sempre confiado em mim. Queria agradecer por fim ao teu companheirismo e a todos os momentos bons que passamos.

Aos restantes membros do laboratório do grupo Macromolecular Structure e Biomolecular Structure and Function, obrigada por tornarem este ano melhor.

Queria agradecer à minha família (à minha mãe, avó e pai) por estarem sempre presentes, mesmo a 350 km de distância. Por estarem sempre interessados no meu trabalho, mesmo não percebendo nada, por me acalmarem quando estava menos bem e principalmente por terem muita paciência comigo.

Por fim, queria agradecer aos amigos que sem eles isto não teria sido nada giro. Catarina, Martinho, Madalena e Rita (Tea das cinco), obrigada por continuarem comigo e estarem sempre presentes. Obrigada por fazerem um esforço em ouvir os meus *podcast* de áudios a falar sobre nada e de conseguirem sempre fazer-me sentir melhor.

Queria agradecer ao Gon, ao meu companheiro de viagem, por ter vindo comigo para o Porto e por termos partilhado esta experiência de vivermos sozinhos. Sem ti estes dois anos teriam sido difíceis, mesmo que tu nunca entendas o telemóvel e fiques a dormir até às 15h da tarde. Por favor, volta para o i3s que sinto a tua falta.

Inês, Marta A, Marta M, Rita e Salvas fico muito feliz por ter-vos conhecido nestes dois anos e obrigada por fazerem parte desta minha jornada. Sem dúvida nunca me irei esquecer de todos os lanches e áudios motivadores.

Resumo

Nesta tese, seguimos o desenvolvimento de dois projetos distintos mas interligados, cada um centrado no complexo mundo das enzimas. Embora estes projetos sejam independentes nos seus objetivos específicos, convergem para o estudo da estrutura e função das enzimas.

Capítulo 1: A combustão de combustíveis fósseis emite óxidos de enxofre nocivos para o ambiente, acarretando riscos significativos para a saúde humana e para o ecossistema. Biodessulfurização é uma alternativa mais económica e amiga do ambiente aos métodos convencionais, para remover enxofre do petróleo bruto. Foi demonstrado que o microrganismo *Rhodococcus erythropolis* IGTS8 remove o enxofre de compostos orgânicos através de uma via conhecida como “via 4S”, que envolve 4 enzimas, DszA, DszB, DszC e DszD. A DszD, uma FMN-NADH oxidoreductase, catalisa a redução de FMN a FMNH₂, o passo limitante desta via, fornecendo FMNH₂ a duas monooxigenases da via. O objetivo deste trabalho foi melhorar a eficiência catalítica da DszD através de mutagénese racional de resíduos importantes para a sua atividade, seguida da caracterização bioquímica e estrutural da proteína nativa e das variantes (T62D, T62R, T62L e T62Y). Nós expressámos e purificámos as quatro variantes e comparámos a sua atividade com a da proteína nativa. Nenhuma das variantes estudadas é mais eficiente do ponto de vista catalítico do que a nativa e apesar de termos conseguido obter cristais nos ensaios de cristalização, concluímos que estes não eram cristais de DszD, mas sim de sal. Uma análise efetuada com AlphaFold mostrou a presença de um segmento N-terminal altamente flexível e provavelmente desordenado. Esta região pode dificultar a cristalização desta enzima e, para testar esta hipótese, gerámos uma versão truncada da mesma que vai ser usada nos ensaios de cristalização subsequentes. A determinação da estrutura 3D da DszD permitirá prever com mais confiança substituições benéficas, bem como interpretar os nossos resultados bioquímicos.

Capítulo 2: Interações proteína-proteína estão no centro de muitos processos biológicos. Estas interações estão envolvidas na regulação de vias bioquímicas, cascatas de sinalização e desempenham um papel fundamental no reconhecimento de substratos. Embora os parceiros intervenientes sejam muitas vezes identificados por métodos biofísicos, os pormenores moleculares deste tipo de interação são muitas vezes elusivos e difíceis de obter através de métodos convencionais. A cristalografia em série resolvida no tempo (SSX) permite a determinação da estrutura de moléculas

biológicas ao nível atómico, bem como o estudo da dinâmica das interações proteína-proteína, do reconhecimento de substratos e do processamento enzimático. Com este trabalho, pretendemos otimizar protocolos para a preparação de microcristais de enzimas (por exemplo, IscS, tripsina e trombina humana), bem como a preparação de microcristais em complexo com inibidores e substratos fotoativáveis, para revelar os detalhes moleculares do reconhecimento do substrato e do ciclo de hidrólise. Neste projeto, desenvolvemos protocolos de *microseeding* para produzir microcristais homogéneos adequados para SSX na linha ID29 do ESRF (Grenoble, França), o qual é uma fonte de luz de sincrotrão de 4ª geração. O objetivo é depois aplicar este protocolo a outras proteínas para obter microcristais para SSX. Após duas fases de *seeding* foi possível produzir microcristais de IscS de tamanho homogéneo adequados para SSX, utilizando o método de difusão de vapor.

Palavras-chave: BDS, DszD, engenharia de proteínas, catálise enzimática, IscS, cristalografia de raios X, SSX, microcristais

Abstract

In this thesis, we follow the story of two distinct yet interconnected projects, each one centered on the intricate world of enzymes. Although these projects are independent in their specific aims, they converge into structural and functional studies of enzymes.

Chapter 1: Combustion of fossil fuels emits harmful sulfur oxides into the environment, posing significant risks to human health and to the ecosystem. Biodesulfurization is a more economical and environmentally friendly alternative to conventional methods such as HDS for removing sulfur from crude oil. The microorganism *Rhodococcus erythropolis* IGTS8 has been shown to remove sulfur from organic compounds through a pathway known as the “4S pathway”, which involves 4 enzymes, DszA, DszB, DszC, and DszD. DszD, an FMN-NADH oxidoreductase, catalyses the reduction of FMN to FMNH₂, the rate limiting step this pathway, supplying FMNH₂ to two monooxygenases in the pathway. The aim of this work was to improve the catalytic power of DszD through rational mutagenesis of key residues of its activity, followed by biochemical and structural characterization of the *wildtype* protein and of the engineered variants (T62D, T62R, T62L and T62Y). We expressed and purified the four variants and compared their activity to the *wt* protein. None of the variants studied is more catalytically efficient than the *wt* protein and, although we were able to obtain crystals in the crystallization screenings, we concluded that these were not DszD crystals, but salt crystals. An analysis performed with AlphaFold showed the presence of a highly flexible and disordered N-terminal region in DszD. This region may be hindering the crystallization of this enzyme and affecting the determination of the apparent molecular weight for each oligomerization state of DszD as FMN concentrations increases. In order facilitate the crystallization of DszD, we generated a truncated version of this enzyme. The determination of the 3D structure of DszD could allow us to predict beneficial substitutions, as well as interpret our biochemical results.

Chapter 2: Protein-protein interactions are at the core of many biological processes. Although the intervening partners are often identified by biophysical methods, the molecular details of this type of interaction are often elusive and difficult to obtain using conventional methods. Time-resolved serial crystallography allows the determination of macromolecular structures at the atomic level, as well as the study of the dynamics of protein-protein interactions, substrate recognition and enzymatic processing. With this work, we aim to optimize protocols for the preparation of microcrystals of enzymes (e.g., IscS, trypsin and human thrombin), as well as the preparation of microcrystals in complex

with inhibitors and photoactivatable substrates, to reveal the molecular details of substrate recognition and the hydrolysis cycle. In this work, we design a straightforward approach using microseeding to produce homogeneously sized microcrystals suitable for SSX in the ID29 beamline at ESRF (Grenoble, France) using 4th generation synchrotron light source. After two-step seeding process, we produced homogeneously sized microcrystals of IscS using the vapour-diffusion method.

Keywords: BDS, DszD, enzymatic catalysis, protein engineering, IscS, X-ray crystallography, SSX, microcrystals.

Table of Contents

List of Tables	x
List of Figures	xi
List of Abbreviations	xiii
Chapter 1: Biochemical and structural characterization of DszD, a desulfurization enzyme of the biodesulfurization 4S pathway	1
INTRODUCTION	2
1.1. Desulfurization methods	3
1.1.1. Biodesulfurization	5
1.2. Biodesulfurization pathways	6
1.2.1. Anaerobic sulfur specific pathway	7
1.2.2. Aerobic sulfur specific pathway: Kodama pathway	7
1.2.3. Aerobic sulfur specific pathway: 4S pathway	7
1.3. DszD, a flavin reductase of the 4S pathway	10
1.3.1. Enzyme characterization and mutagenesis studies	11
Motivation and Aim	15
MATERIALS AND METHODS	17
2.1. Materials	18
2.1.1. Bacterial Strains	18
2.1.2. Cell Culture Media	18
2.1.3 Buffer Solutions	19
2.2. Methods	19
2.2.1. Single-site mutagenesis	19
2.2.2. Plasmid DNA Replication	20
2.2.3. Protein Expression	20
2.2.4. Protein Purification	21
2.2.5. SDS-Page analysis	22
2.2.6. Agarose gel electrophoresis of DNA analysis	22

2.2.7. Activity assays	22
2.2.8. Size Exclusion Chromatography	23
2.2.9. Crystallization Screenings	23
RESULTS	25
3.1. Expression and purification of DszD <i>wild type</i> and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y)	26
3.2. Activity assays of DszD <i>wt</i> and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y)	29
3.3. Effect of FMN on DszD oligomerization	34
3.4. DszD's AlphaFold structure prediction	36
3.5. Crystallization of DszD <i>wt</i>	38
DISCUSSION	41
Conclusions and Future Perspectives	45
Chapter 2: Application of time-resolved serial crystallography on a 4th generation synchrotron light source to dissect the mechanisms of enzyme inhibition and activity at room temperature	47
INTRODUCTION	48
1.1. Protein Structure Determination	49
1.1.1. Protein crystallization	49
1.2. X-ray Crystallography	53
1.3. Time-resolved serial crystallography	53
1.3.1. Sample requirements	56
1.3.2. Obtaining microcrystals	56
1.3.3. Sample delivery	56
1.3.4. Reaction initiation	58
1.3.5. Importance of optimizing protocols for time-resolved serial crystallography	59
1.3.5.1 Cysteine desulfurase, IscS	60
Motivation and Aim	62
.....	63
MATERIALS AND METHODS	63

2.1. Materials	64
2.1.1. Bacterial Strains	64
2.1.2. Cell Culture Media	64
2.1.3 Buffer Solutions	65
2.2. Methods	65
2.2.1. DNA replication.....	65
2.2.2. Protein Expression	66
2.2.3. Protein Purification	66
2.2.4. SDS-PAGE analysis	67
2.2.5. Crystallization	67
RESULTS	69
3.1. Expression and purification of IscS	70
3.2. IscS crystallization for time-resolved serial crystallography	72
DISCUSSION	76
Conclusions and Future Perspectives	79
References	81
Annexes.....	90

List of Tables

Table 1. Gibbs activation and reaction free energies for the hybrid transfer in each of the variants tested.	12
Table 2. Bacterial strains used in the following protocols.....	18
Table 3. Cell culture media used in the following protocols.....	18
Table 4. Buffers used in the following protocols.	19
Table 5. Primers used for single-site mutagenesis.....	20
Table 6. Kinetics parameters (k_{cat} , K_m and V_{max}) of DszD <i>wt</i> and variants, DszD_T62D, DszD_T62R, DszD_T62L and DszD_T62Y.....	34
Table 7. K_{av} and the predicted molecular weight of DszD <i>wt</i>	36
Table 8. Primers used in IVA cloning.	37
Table 9. K_{av} and the predicted molecular weight of ΔNT -DszD..	38
Table 10. Bacterial strains used in the following protocols.....	64
Table 11. Cell culture media used in the following protocols.....	64
Table 12. Buffers used in the following protocols.	65
Table 13. Composition of the crystallization conditions used from Morpheus® MD1-46 (Molecular Dimensions) conditions.....	90

List of Figures

Figure 1. Anaerobic conversion of DBT to biphenyl.	7
Figure 2. Scheme of the aerobic pathways for DBT degradation.	9
Figure 3. The two-step conversion of FMN to FMNH ₂ catalysed by DszD..	10
Figure 4. Predicted three-dimensional structure of the DszD enzyme..	12
Figure 5. DszD's activity assay set-up.	23
Figure 6. DszD <i>wt</i> protein expression.	26
Figure 7. DszD <i>wt</i> purification by Immobilized Metal Affinity Chromatography..	28
Figure 8. SDS-PAGE gel of DszD <i>wt</i> and variants purified using two IMAC purification steps.	29
Figure 9. Linear dependence of the initial velocity with DszD <i>wt</i> enzyme concentration, at different NADH concentrations..	30
Figure 10. NADH and NAD ⁺ concentration curve measured using the absorbance at 340 nm.	31
Figure 11. Consumption of NADH by DszD <i>wt</i> and variants DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y as a function of time (300 s).	32
Figure 12. The impact of substitutions of residue 62 in the enzymatic activity of DszD.	33
Figure 13. The effect of FMN on the oligomerization of DszD <i>wt</i>	36
Figure 14. DszD structure predicted by AlphaFold merged with Sousa et al. (2016) model.	37
Figure 15. SEC elution profiles of ΔNT-DszD and FL-DszD.	38
Figure 16. Crystals obtained from two Morpheus® MD1-46 (Molecular Dimensions) conditions.	39
Figure 17. SEC chromatogram of DszD <i>wt</i> in the presence of FMN.	40
Figure 18. Crystals obtained from B11 condition of Crystal Screen Cryo screen (Hampton Research).	40
Figure 19. Asymmetric unit, unit cell and entire crystal.	50
Figure 20. Phase diagram illustrating the different routes to reach supersaturation. . .	50
Figure 21. Vapor diffusion method.	52
Figure 22. Experimental setup for soaking protein crystals in a sitting drop.	53
Figure 23. Schematic illustration of a TR-SSX.	55
Figure 24. Schematic illustration of sample delivery methods for TR-SSX..	57
Figure 25. Schematic diagram of a photocaging reaction..	59
Figure 26. Simplified illustration of the enzymatic mechanism of IscS..	60

Figure 27. S-(2-nitrophenyl)-L-Cysteine structure.	61
Figure 28. Seed preparation scheme using Seed Bead™ (Hampton Research protocol)..	68
Figure 29. SDS-PAGE gel of IscS protein expression.....	70
Figure 30. SDS-PGE gels of IscS after IMAC purification..	72
Figure 31. Analysis of IscS purification by SEC..	72
Figure 32. IscS crystals in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions)..	73
Figure 33. IscS crystals in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions).	74
Figure 34 IscS crystals grown in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions) with 1 st round seeds.....	74
Figure 35. IscS crystals grown in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions) with 2 nd round seeds..	75
Figure 36. Map of the plasmid used for recombination expression of DszD <i>wt</i> and variants (T62D, T62R, T62L, and T62Y).	90

List of Abbreviations

AcN – Acetonitrile

ADS – Adsorptive desulfurization

BDS – Biodesulfurization

DBT – Dibenzothiophenes

CAM – Chloramphenicol

CV – Column Volume

DBTO – DBT sulfoxide

DBTO₂ – DBT sulfone

(ΔN-DszD) – N-terminal truncated version of DszD

DMSF – Dimethylformamide

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic Acid

DTT – Dithiothreitol

EDS – Extractive desulfurization

ESRF-EBS – European Synchrotron Radiation Facility - Extremely Brilliant Source

FID – Free interface diffusion

FL-DszD – Full-length DszD

FMN – Flavin mononucleotide

FMNH₂ – Reduced flavin mononucleotide

FT – Flow-through

Glu – Glutamic acid

2HBT – 2-hydroxybiphenyl

HDS – Hydrodesulfurization

HpaCst – -Hydroxyphenylacetate hydroxylase component C

IPTG – Isopropyl β -D-1-thiogalactopyranoside

IMAC – Immobilized metal affinity chromatography

IVA – *In vivo* Assembly cloning

KAN – Kanamycin

LB – Lysogeny Broth

MW – Molecular weight

MR – Molecular replacement

NADH – Nicotinamide adenine dinucleotide

NMR – Nuclear magnetic resonance

NPS – Nitrate phosphate sulfate

OD – Optical Density

ODS – Oxidative desulfurization

PEG – Polyethylene Glycol

PCR – Polymerase Chain Reaction

QM/MM – Quantum mechanics/molecular mechanics

Rpm – Revolutions per minute

SAD/MAD – Single-wavelength anomalous diffraction/multi-wavelength anomalous diffraction

SDS-PAGE – Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SEC – Size Exclusion Chromatography

SO₃ – Sulfite

SO_x – Sulfur oxides

SSX – Serial Synchrotron Crystallography

TAE – Tris-Acetate-EDTA

TEV – Tobacco Etch Virus Thr (T) – Threonine

Trx – Thioredoxin

TR-SSX – Time-resolved Serial Synchrotron Crystallography

wt – wild type

XFEL – X-ray free-electron laser

V_0 – Initial velocity

V_0 – Elution Volume

$V_{\max}$ – Maximum velocity

Chapter 1: Biochemical and structural characterization of DszD, a desulfurization enzyme of the biodesulfurization 4S pathway

INTRODUCTION

*Biochemical and structural characterization of DszD, a desulfurization enzyme of the
biodisulfurization 4S pathway*

Combustion of fossil fuels is currently one of the main sources of environmental pollution releasing into the atmosphere various hazardous pollutants. The emission of these pollutants, such as sulfur oxides (SO_x) can lead to air pollution, acid rain, and subsequent human health issues, such as respiratory problems (Mohebbali & Ball, 2008). Part of this problem is due to the heavy presence of sulfur in crude oil, being the most abundant element in petroleum after carbon and hydrogen with amounts ranging from 0.1% to 6% (w/w), and exceeding 14% in some crude oils and asphalts (Shi & Wu, 2021). Due to its harmfulness, several governments have imposed stricter restrictions on the sulfur content in crude oil, and automakers and refiners must meet these legal requirements for vehicles and fuels production (Ferreira *et al.*, 2017). It is therefore necessary to reduce the sulfur content of fossil fuels to reduce sulfur oxides emissions.

Crude oils consist of a mixture of several different hydrocarbon molecules with different carbon numbers ranging from $\text{C}_1\text{-C}_4$ to $\text{C}_{12}\text{-C}_{70}$ (Demirbas & Bamufleh, 2017). During the refinement process of crude oil, the concentration of sulfur in these compounds changes according to the boiling range and molecular weight, increasing with the higher molecular weight components of the crude oil (Mohebbali & Ball, 2008). Sulfur compounds are present in two forms, inorganic and organic (e.g., thiols, sulfides, thiophenic), with the latter being the major source of sulfur in crude oil, while more complex sulfur compounds are typically found in crude oils with higher viscosities and densities. The nature of these compounds can also affect the removal of sulfur from fossil fuels; for example, aliphatic acyclic sulfides are much easier to remove than compounds where sulfur is contained within aromatic rings, such as alkylated dibenzothiophenes (DBT), which represent the largest fraction of sulfur compounds in crude oils (Javadli & De Klerk, 2012). Due to the need to improve the quality of fossil fuels and to decrease their sulfur content, the development of methods targeting the desulfurization of fossil fuels have been the subject of intense research in the past decades.

1.1. Desulfurization methods

Hydrodesulfurization

Hydrodesulfurization (HDS) is the most employed method by the petroleum industry to extract sulfur from fossil fuels. In HDS, the crude oil is fed into a fixed-bed reactor containing a catalyst (e.g., $\text{Co-Mo/Al}_2\text{O}_3$ or $\text{Ni-Mo/Al}_2\text{O}_3$) with hydrogen (H_2), which is converted to hydrogen sulfide (H_2S) and further removed from the oils. This process

requires harsh operating conditions with high temperatures (~ 400 °C) and pressures (~ 20 MPa) to achieve the desired low level of sulfur. This method removes relatively simple sulfur compounds, such as aliphatic sulfur compounds, and fails to remove more complex aromatic sulfur compounds such as alkylated dibenzothiophenes and their derivatives due to their length and alkyl positions. After HDS, these compounds form the most abundant sulfur compounds in the refined oil (Ma, 2010).

Due to the harsh operating conditions, which decreases the energetic value of the fuel (Abro *et al.*, 2022), and the failure to remove refractory sulfur compounds, alternative methods to HDS are being investigated to obtain fuel with low sulfur levels that comply with sulfur regulations. These alternative approaches for sulfur removal are (i) oxidative desulfurization (ODS); (ii) extractive desulfurization (EDS); (iii) adsorptive desulfurization (ADS); and (iv) biodesulfurization (BDS) (Boniek *et al.*, 2015).

Oxidative desulfurization

Oxidative desulfurization (ODS) is characterized by a reaction between an oxidizing agent and sulfur to facilitate its removal from the crude oil. In addition, ODS can be used as a complementary process to HDS to remove refractory sulfur compounds that are not removed from crude oils by HDS, particularly aromatic organosulfur compounds like benzothiophene, dibenzothiophene and derivatives (Rajendran *et al.*, 2020). The ODS method consists of two stages: (i) the selective oxidation of the aromatic sulfur compounds resulting in their conversion into sulfones and sulfoxides, and (ii) their extraction by liquid-liquid extraction, adsorption, or distillation due to changes in their polarity (Gangwal, 2011). Sulfones and sulfoxides are more polar compared to other hydrocarbons present in the fuel, allowing them to be easily removed from the oil.

ODS is an attractive method to HDS because it is more energetically efficient, operating under mild temperatures (25-90 °C) and pressures (0.1 MPa), and it is more environmentally friendly since it uses an eco-friendly oxidizing agent, such as H_2O_2 (Gangwal, 2011). However, this method has its limitations, such as the removal of additional compounds of interest alongside sulfones and sulfoxides, waste management of oxidized sulfur compounds, as well as the recovery and reuse of oxidants and catalysts (Saha *et al.*, 2021).

Extractive desulfurization

Extractive desulfurization (EDS) is a method that extracts organosulfur compounds from crude oils without affecting other compounds in the oil (Jha *et al.*, 2023). EDS is based on the solubility difference between organosulfur compounds in certain extractant

solvents. In this method, the fossil fuel is mixed with the extract solvent, forming two immiscible phases. After the mixing, the organosulfur compounds are extracted into the solvent due to their higher solubility in the solvent (Javadli & De Klerk, 2012). EDS often uses dimethyl sulfoxide (DMSO), dimethylformamide (DMF), pyrrolidones, and acetonitrile (AcN) as extractant solvents (Saha *et al.*, 2021).

EDS is an appealing method because it uses mild operating conditions (ambient temperatures and pressure), provides straightforward separation of the refined fuel and is inexpensive compared to other methods. Furthermore, EDS is a non-destructive method that does not alter the quality of the fossil fuel (Moghadam *et al.*, 2020). However, experimental findings showed that using the above organic solvents to remove sulfur is very inefficient. Additionally, this method leads to a substantial loss in fuel volume due to co-extraction of a significant amount of aromatic hydrocarbons along with the organosulfur compounds, as a result of their similar polarities (Saha *et al.*, 2021).

Adsorptive desulfurization

Adsorptive desulfurization (ADS) consists of the removal of organic sulfur compounds from crude oil by physicochemical adsorption processes (Saha *et al.*, 2021). In ADS, sulfur compounds interact with the adsorbent and become retained on its surface by both physical and chemical interactions. The efficiency of this process depends on the characteristics of the adsorbents, such as high surface area and pore volume, more mesopores and surface-active sites, and good structural strength and stability. Furthermore, the type of adsorbent used is important for the efficiency of this process, and different adsorbents have been used, such as activated carbon, modified carbon, mesoporous material, zeolites, metal oxide, metal-organic frameworks (Lee & Valla, 2019).

ADS is a good alternative method to HDS, because it requires mild temperatures and pressures (low energy demand), it is easy to operate, and it is low cost due to the relatively cheap adsorbent. However, the main challenge in this process is the regeneration of the adsorbent, which can be difficult to achieve without any loss of adsorption capacity (Omar & Verma, 2022).

1.1.1. Biodesulfurization

Biodissulfurization is one of the most innovative approaches for replacing the above methods. As the name suggests, this method uses microorganisms to specifically remove sulfur atoms from organic sulfur compounds without compromising the carbon skeleton of the compounds. These microorganisms possess enzymes that selectively

oxidize sulfur by cleaving carbon-sulfur bonds, converting DBT and its alkylated derivatives to non-sulfur compounds and sulfite (Mamuad & Choi, 2023).

The idea of using microorganisms to remove sulfur from crude oil first came to light in 1935, and in the early 1950s several patents were issued for the application of microorganisms in the petroleum industry to desulfurize fossil fuels. However, BDS was quickly dismissed because the isolated bacteria were not suitable for commercial desulfurization (Ohshiro & Izumi, 1999). Since then, research on the application of these microorganisms in the petroleum industry has accelerated, and according to Sadare *et al.* (2017), several studies have been conducted on different types of microorganisms for applications in the refining process, both at the batch scale and at the industrial scale. Some of these applications are biological demulsification, denitrogenation, demetallation, enhancement of oil recovery by microorganism, conversion of heavy crude to light crude, and biological control of reservoir souring (Sadare *et al.*, 2017). Although the large-scale application of the BDS technology is still in its early stages, the U.S. Company Energy Bio-Systems Corporation has developed a pilot project where sulfur can be removed from up to 5 barrels of oil per day. In this project, the biocatalyst is mixed with the crude oil in a bioreactor where the biological desulfurization takes place (Boniek *et al.*, 2015). According to Monticello (2000), this project marks the beginning of the industrial process of BDS (Monticello, 2000).

The main advantage of BDS is being environmentally friendly, as it operates at ambient temperatures and atmospheric pressures (less energetically demanding), and does not use chemicals, as the aforementioned methods (Stylianou *et al.*, 2021). In addition, BDS removes refractory sulfur compounds that HDS cannot remove, like DBT and alkylated derivatives, at mild operating conditions. Due to this, BDS can be used as an alternative to HDS or a complementary process to remove sulfur from fossil fuel. Other great advantage of BDS is that greenhouse emissions can be significantly reduced, in comparison to HDS (Sadare *et al.*, 2017). BDS proves to be a better alternative to HDS and to the other methods, because these other degradation pathways rely on the oxidation and mineralization of the carbon skeleton instead of sulfur removal, significantly reducing the fuel value of the end product (Grossman *et al.*, 1999).

1.2. Biodesulfurization pathways

Desulfurizing microorganisms require sulfur, which makes up 0.5 to 1% of their dry weight, for growth and essential physiological activities. In nature, sulfur is typically found in organic forms, which are less readily available than inorganic sulfates. As a result, bacterial cells have developed mechanisms to extract sulfur from these organic

compounds. To date, three main pathways for biodesulfurization have been reported, including one anaerobic pathway and two aerobic pathways (the Kodama pathway and the 4S pathway).

1.2.1. Anaerobic sulfur specific pathway

Sulfate-reducing bacteria have been reported to desulfurize model compounds and fossil fuels under anaerobic conditions by Kim *et al.* (1995). *Desulfovibrio desulfuricans* M6 is able to selectively remove sulfur by converting DBT to biphenyl as the major reaction product and H₂S (**Figure 1**) (Kim *et al.*, 1995). Desulfurization of fossil fuels under anaerobic conditions eliminates the cost of aeration and has the advantage of releasing sulfur as a gas. However, the development of an anaerobic BDS process is currently hindered by low reaction rates, safety issues and cost concerns (Mohebbi & Ball, 2016).

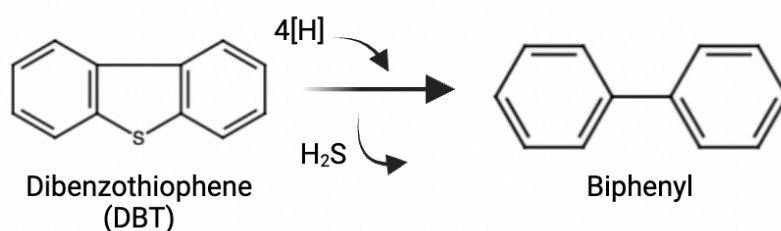


Figure 1. Anaerobic conversion of DBT to biphenyl. In the anaerobic pathway, DBT is reduced to biphenyl and H₂S. Scheme adapted from (Stylianou *et al.*, 2021).

1.2.2. Aerobic sulfur specific pathway: Kodama pathway

The Kodama pathway is a destructive ring pathway in which the C-C bond of one of the benzene rings of DBT is selectively cleaved through a series of oxidation steps, leaving the C-S bond in the partially cleaved benzyl ring intact. In this pathway, DBT is degraded to a water-soluble organosulfur compound and pyruvate, and the sulfur remains in the reaction products (Kalita *et al.*, 2022). In the Kodama pathway, the degradation process of DBT begins with its oxidation by dioxygenases, then DBT is degraded by ring opening (**Figure 2**). In this pathway no desulfurization of the organosulfur substrate occurs, since the C-S bond is retained and converted to small molecular organic sulfides (L. Wang *et al.*, 2019).

1.2.3. Aerobic sulfur specific pathway: 4S pathway

The 4S pathway is a sulfur-specific pathway in which only the C-S bond in DBT is cleaved, leaving the carbon skeleton of the compound intact. This pathway was found in the bacterium *Rhodococcus erythropolis* IGTS8, and this organism has been shown to be able to use DBT as a source of sulfur. The end product of this pathway is 2-hydroxybiphenyl (2HBT), and the sulfur is released in the form of sulfite (SO₃²⁻),

preserving the calorific value of the fossil fuel. This does not happen in the Kodama pathway, which makes the 4S pathway a promising technology for fossil fuel biodesulfurization. The pathway was named “4S” due to the four intermediates formed: DBT sulfoxide (DBTO); DBT sulfone (DBTO₂); 2HBT; and SO₃ (Kalita *et al.*, 2022).

The 4S pathway is divided into four oxidation reactions to completely remove sulfur from DBT. These reactions are catalysed by a multienzymatic complex encoded by the *dsz* genes, all working under the *dszABC* operon of *R. erythropolis* IGTS8. This pathway starts with two consecutive oxidation reactions of DBT to produce DBTO₂, both catalysed by DszC (DBT monooxygenase), a step that has been shown to be slow. Then, the C-S bond of the thiophane ring is cleaved by DszA (DBTO₂ monooxygenase). Finally, the sulfur is removed resulting in the formation of 2HBT and SO₃, catalyzed by DszB (desulfinate enzyme) (**Figure 2**) (Gray *et al.*, 1996). In addition, both DszC and DszA enzymes require a supplementary enzyme that supplies them with reduced flavin mononucleotide (FMNH₂), essential for their activity. This key enzyme is DszD, a flavin reductase that recovers the cofactor FMNH₂ by oxidizing reduced β-Nicotinamide adenine dinucleotide (NADH) for the first three oxidation reactions (Soleimani *et al.*, 2007). The final products of the 4S pathway, 2HBT and SO₃⁻, have an inhibitory effect on the desulfurization reaction. It has been reported that the accumulation of sulfite is toxic for the growth of the bacteria, as they cannot use the sulfur from the sulfite. Furthermore, the activity of both DszA and DszB is inhibited by 2HBT (Chen *et al.*, 2018).

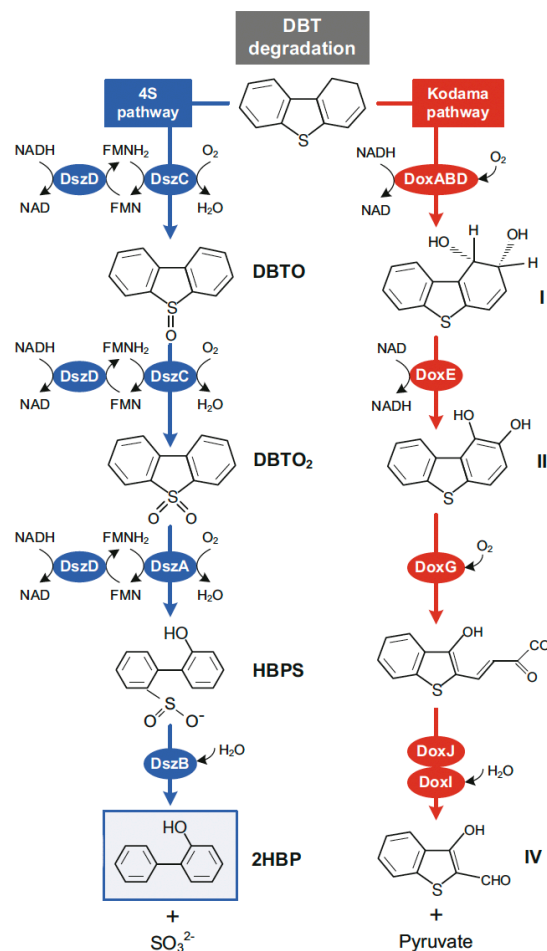


Figure 2. Scheme of the aerobic pathways for DBT degradation. The 4S pathway is represented in blue and the Kodama pathway represented in red. The metabolites shown are DBT dibenzothiophene; DBTO dibenzothiophene sulfoxide; DBTO₂ dibenzothiophene sulfone; 2-hydroxybiphenyl-2-sulfinate (HBPS); 2-hydroxybiphenyl (2HBP); 2-methoxybiphenyl; (BP) biphenyl; I cis-1,2-dihydroxy-1,2-dihydroDBT; II 1,2-dihydroxyDBT; III cis-4-[2-(3-hydroxy)-thionaphthenyl]-2-oxo-3-butenate; IV 3-hydroxy-2-formyl-benzothiophene. Scheme adapted from (Martínez *et al.*, 2017).

Several advances in BDS have been made over the decades to improve the quality and reduce the sulfur content of fossil fuels. Much of this attention has focused in the microbiological and biochemical methods to desulfurize crude oil. The motivation for this work arose from the discovery of a highly specific metabolic pathway in *R. erythropolis* IGTS8, the 4S pathway. This organism can extract nutritional sulfur from heterocyclic molecules in fuel, like DBT, without causing oxidative carbon loss from the oil. Although this method is a promising alternative to the traditional desulfurization methods, there are several factors that prevent its adoption in the industry. One of the main issues behind the failure of BDS in biorefineries is the relatively slow reaction rate of the biodesulfurization process in the 4S pathway, which is much slower than the chemical methods, like HDS. One way to overcome this limitation is to improve the catalytic power

of the enzymes involved in this reaction through rational mutagenesis of key residues (Ferreira *et al.*, 2017).

1.3. DszD, a flavin reductase of the 4S pathway

DszD is a flavin reductase identified in *R. erythropolis* IGTS8 that supplies FMNH₂ to DszC and DszA in the 4S pathway (**Figure 2**). DszD is a 192 amino acid protein, with a predicted molecular weight of 20 kDa (Kamali *et al.*, 2010), encoded by the *dszD* gene located on the IGTS8 chromosome, rather than on the plasmid containing the *dszABC* operon (Kalita *et al.*, 2022). In the 4S pathway, DszD catalyses the reduction of flavin mononucleotide (FMN) to FMNH₂ through the oxidation of NADH to NAD⁺. In this reaction, DszD transfers the nicotinamide moiety of NADH to the central nitrogen of the oxidized FMN, forming a semiquinone intermediate (FMNH) (**Figure 3**, step 1). Then, a proton transfer occurs to FMNH, resulting in the reduction of FMN (**Figure 3**, step 2). Gray *et al* (2003) and other authors propose that this enzyme may be responsible for the rate-limiting step of the 4S pathway (Gray *et al.*, 2003).

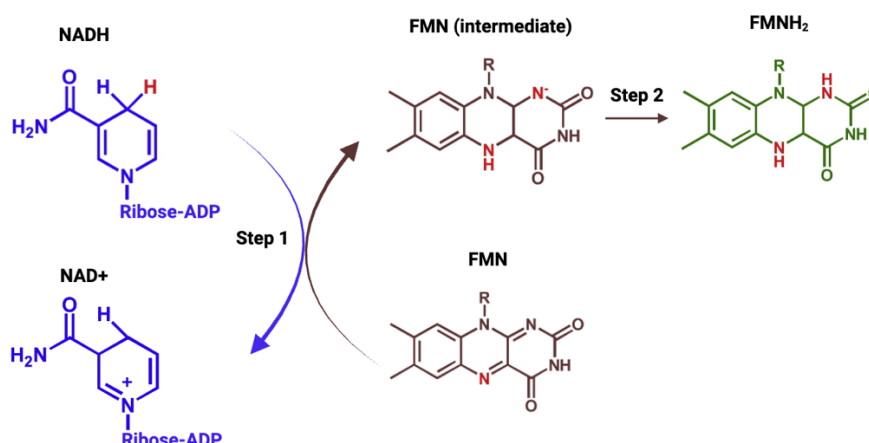


Figure 3. The two-step conversion of FMN to FMNH₂ catalysed by DszD. The first step is represented by the transfer of the nicotinamide moiety of NADH to the central nitrogen in the isoalloxazine moiety of the oxidized FMN, forming a semiquinone intermediate (FMNH). In the second step, occurs a proton transfer to FMNH, leading to the reduction of FMN. Scheme adapted from (Li *et al.*, 2020).

One strategy proposed to overcome the rate at which the bacterial cells remove sulfur in BDS is to increase the enzymatic activity of the DszD enzyme through rational mutagenesis of important residues for its activity, as this enzyme supplies FMNH₂ to DszA and DszC, and is responsible for the rate-limiting step of BDS. The enhancement of the DszD activity should contribute to a faster desulfurization of crude oil (Sousa *et al.*, 2016).

1.3.1. Enzyme characterization and mutagenesis studies

Currently, there is no X-ray or NMR structure of the DszD enzyme from *R. erythropolis* IGTS8, either in its apo-form or complexed with both NADH and FMN. However, according to Kamali *et al* (2010), the crystal structure of 4-hydroxyphenylacetate hydroxylase component C from *Sulfolobus tokodaii* strain 7 (HpaCst) has the closest structural homology to the *R. erythropolis* IGTS8 DszD enzyme, with a highly superimposable active site and very close steric orientation of the active site substrates (Kamali *et al.*, 2010). Moreover, in this structure, both substrates, FMN and NADH, are already at an adequate distance for the step 1 of the reaction to occur (**Figure 3**). Hence, the HpaCst structure was used as template to predict the three-dimensional structure of *wild type* (*wt*) DszD. In addition, this model also predicts residues that directly interact with both FMN and NADH, enabling the identification of a key residue at position 62 of the polypeptide that is thought to be directly involved in the rate-limiting step of the reduction of FMN to FMNH₂. This residue is located near (within 3 to 3.5 Å) the N5 atom of the isoalloxazine ring of FMN (**Figure 4**) (Sousa *et al.*, 2016). Previous mutagenesis studies done by Kamali and coworkers demonstrated that replacing DszD Thr62 with asparagine (T62N) or alanine (T62A) increased the catalytic rate by 5 and 7-fold, respectively. It has been suggested that the nature of this residue in step 1 is to regulate the reaction by slowing down its rate, rather than accelerating it (Kamali *et al.*, 2010).

To further investigate the catalytic mechanism of DszD, Sousa *et al* (2016) examined the atomistic and energetic profile of the mechanism of DszD using hybrid quantum mechanics/molecular mechanics (QM/MM) methods. In addition, this study suggests strategies for improving the biocatalytic desulfurization of crude oil. The study was conducted using the homologous enzyme HpaCst of *S. tokodaii* strain 7 as model. However, there is one subtle difference between the sequence of DszD and HpaCst, an active site residue that acts as close spectator, which in HpaCst is an asparagine located at position 33 and in DszD is a threonine located at position 62. Due to the high similarity of both active sites, the conclusions are thought to be transferable (Sousa *et al.*, 2016).

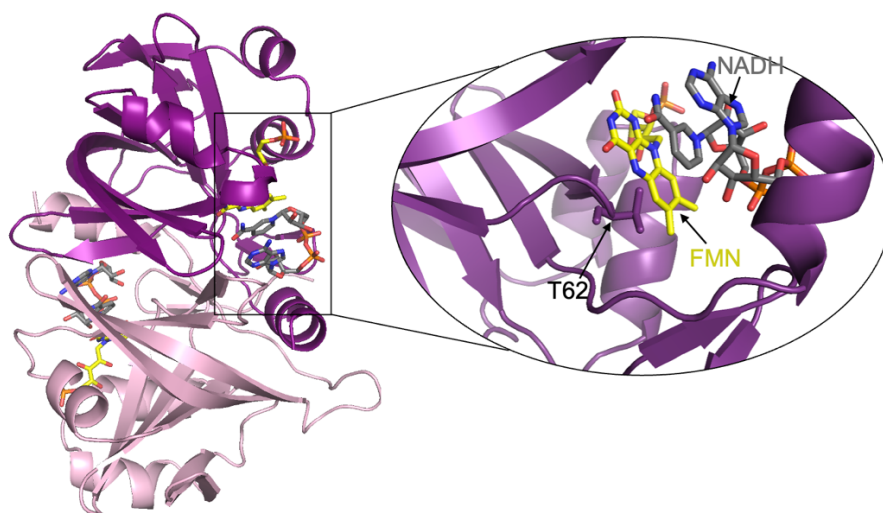


Figure 4. Predicted three-dimensional structure of the DszD enzyme. Cartoon representation of the homodimer of DszD. The two monomers are colored purple and pink in cartoon representation. The FMN (yellow) and NADH (grey) molecules are represented as sticks. Structural model from Sousa *et al.* (2016).

In this study, Sousa and coworkers analyzed the effect of different mutations on the activity of DszD, particularly focusing on the active site Thr62 residue and how to optimize the efficiency of this process. They started by replacing Thr62 with asparagine and alanine by computational single site-directed mutagenesis. The results showed that these substitutions led to a decrease in the activation barrier of the first of the two steps of the reduction of FMN to FMNH₂, aligning with experimental results obtained by Kamali *et al.* (2010). These results suggest that the presence of smaller or positively charged amino acids at this site could lower the barrier height of the rate limiting steps. These findings can be applied in rational protein engineering of new DszD variants to improve the rate at which BDS can be performed as this enzyme provides FMNH₂ to two other enzymes in the 4S pathway, DszA and DszC (Sousa *et al.*, 2016).

Sousa *et al.* (2016) and Ferreira *et al.* (2017) tested 18 different amino acids substitutions of Thr62 using computational methods, starting from the model used in the work of Sousa *et al.* (2016), to accelerate the limiting step of the DszD reaction (**Table 1**) (Ferreira *et al.*, 2017). The finding presented in Ferreira *et al.* (2017) were the starting point for our work presented in the following chapters.

Table 1. Gibbs activation and reaction free energies for the hybrid transfer in each of the variants tested. The values were obtained with the enzyme in vacuum, or with solvent, water and heptane, applying the ONIOM-PCM methodology within Gaussian. Table from Ferreira *et al.* (2017).

Variant	$\Delta G^{\ddagger}_{\text{vacuum}}$ [kcal mol ⁻¹]	$\Delta G^0_{\text{vacuum}}$ [kcal mol ⁻¹]	$\Delta G^{\ddagger}_{\text{water}}$ [kcal mol ⁻¹]	$\Delta G^0_{\text{water}}$ [kcal mol ⁻¹]	$\Delta G^{\ddagger}_{\text{heptane}}$ [kcal mol ⁻¹]	$\Delta G^0_{\text{heptane}}$ [kcal mol ⁻¹]
N33D	15.1	-2.3	14.5	-3.4	15	-2.4
N33L	17.8	-0.1	14.8	-4.1	17	-1.2
N33M	18	-3.9	16.9	-3.7	17.8	-3.9
N33HIE	18.3	-3.2	17.7	-4.1	18.2	-3.3
N33Y	18.5	1.3	17.8	0.0	18.3	1.0
N33I	18.5	-0.1	16.4	-2.9	17.8	-0.9
N33S	19.2	1.6	18.6	1.2	19.2	1.6
N33F	19.4	0.8	16.2	-3.0	18.5	-0.1
N33E	19.7	-0.2	18.7	-2.0	19.5	-0.5
N33V	20.6	5.5	21.2	5.3	20.6	5.3
N33HID	22.8	2.5	23.7	3.0	23.1	2.6
N33HIP	23.0	7.8	21.9	6.2	22.7	7.4
N33G	23.2	7.4	23.0	6.8	23.2	7.4
N33Q	23.7	4.8	23.2	2.5	23.5	4.2
N33A	24.6	2.1	23.7	4.4	24.4	2.3
N33W	25.1	8.5	21.4	4.1	24	7.5
WT	25.5	3.3	24.5	2.7	25.1	3.1
N33C	25.6	5.8	25.3	5.4	25.6	5.8
N33T	25.7	2.2	24.9	1.3	25.5	2.0
N33K	26.1	7.5	24.9	5.3	25.9	7.2
N33R	29.9	8.8	29.8	8.2	29.9	8.6

The results of this study showed that substitution of Thr62 with aspartate or glutamate resulted in a lower free energy barrier (15.1 kcal.mol⁻¹ and 19.7 kcal.mol⁻¹, respectively) than the *wt* (25.5 kcal.mol⁻¹). This may be due to their ability to stabilize NAD⁺. Furthermore, positively charged mutants, such as lysine and arginine, displayed an increased activation free energy (26.1 kcal.mol⁻¹ and 29.9 kcal.mol⁻¹, respectively). This may indicate that the positive charge and the bulkiness of lysine is responsible for the high barrier in this mutant and in the arginine mutant, both volume and positive charge may contribute to the reduced reaction rate. Ferreira and colleagues hypothesize that the spectator residue has the function of slowing down the reaction rate to keep the homeostasis of the organism. A higher catalytic rate of DszD would lead to an excess of NAD⁺, which could have several consequences on the homeostasis of the bacteria (Ferreira *et al.*, 2017). To mitigate the impact of this metabolite accumulation and ensure efficient biodissulfurization, further optimization in the bacteria is necessary. One approach is to genetically engineer metabolic pathways that recycle NAD⁺ or convert it back to its reduced form, NADH (Y. Wang *et al.*, 2009). In addition, optimization of culture conditions can help regulate metabolic fluxes and thus effectively manage the levels of NAD⁺ and other metabolites (Nieto-Taype *et al.*, 2020). Bacterial strains used at the industrial scale must be resistant to harsh environments, in which they are exposed to

extreme temperatures and pH, as well as toxic concentrations of metabolites in the large-scale bioreactor (Y. Wang *et al.*, 2009).

Collectively, these studies show that rational protein engineering of new DszD variants can enhance the rate of biodesulfurization, potentially leading to a more efficient biodesulfurization system in *Escherichia coli* and *R. erythropolis*.

Improving the catalytic rate of DszD is one of the key steps in making the 4S pathway attractive for industrial application, as its activity is critical for the activity of the other enzymes involved in the pathway. Increasing the catalytic rate of DszD will increase the yield and the rate of the BDS process, making it a promising biotechnological process capable of reducing the environmental impact and health problems and changing the direction of the crude oil refining industry.

Motivation and Aim

The combustion of fossil fuels emits harmful sulfur oxides into the environment, posing significant risks to human health and to the ecosystem. Currently, there is a significant demand for clean energy sources and a key aspect of this demand is high quality (low sulfur) crude oil. One of the main technologies currently employed to remove sulfur from fossil fuels in refineries is HDS. This method uses harsh conditions (high temperatures and pressures) to remove sulfur from these compounds. However, HDS fails to remove recalcitrant sulfur compounds, such as DBT and derivatives, reducing the calorific value of the crude oil with the harsh operating conditions. One attractive alternative to HDS is BDS that uses microorganisms, such as *R. erythropolis*, to specifically remove sulfur from crude oil without compromising the carbon skeleton of the compounds and preserving the energetic value of the fuel.

The specific sulfur removal process carried out by *R. erythropolis* relies on the 4S pathway, composed by four enzymes: DszA, DszB, DszC, and DszD. DszD is a flavin reductase that supplies FMNH₂ to two monooxygenases in the pathway, DszA and DszC. DszD catalyses the reduction of FMN to FMNH₂ coupled with the oxidation of NADH to NAD⁺, which is the rate-limiting step of the 4S pathway. Hence, increasing the enzymatic activity of the DszD enzyme by rational mutagenesis of residues important for its activity has been proposed as a strategy to increase the rate at which bacteria remove sulfur from crude oil.

The present work aimed at improving the catalytic power of the enzyme DszD through rational mutagenesis of key residues, followed by biochemical and structural characterization of the *wt* protein and its variants. Enhancement of the DszD activity is expected to contribute to faster desulfurization of crude oil. For that purpose, we expressed and purified three variants predicted *in silico* to enhance the catalytic power of DszD and one variant for its predicted negative effect (T62R), and performed the biochemical characterization of these variants, comparing their activity to the *wt* protein. Additionally, we sought to determine the 3D structure of this enzyme using X-ray crystallography to better predict beneficial substitutions, as well as to interpret our biochemical results.

Ultimately, optimizing the catalytic power of DszD in the 4S pathway through rational mutagenesis has the potential to significantly enhance the biodesulfurization process. This could lead to more efficient and cost-effective methods for removing sulfur from

crude oil, addressing an important challenge in the petroleum industry and contributing to cleaner fuel production.

MATERIALS AND METHODS

2.1. Materials

2.1.1. Bacterial Strains

Table 2. Bacterial strains used in the following protocols. List of bacterial strains, their genotype and application in the following protocols.

Bacterial strains	Genotype	Application	Methods section
<i>E. coli</i> DH5 α	<i>F-endA1 glnV44 thi-1 recA1 relA1 gyrA96</i> $\Phi80dlacZ\Delta M15$ <i>deoR</i> $\Delta(lacZYA-argF)$ <i>U169 nupG</i> , <i>hsdR17(rK- mK+)</i> , λ	General cloning and sub-cloning	2.2.2
<i>E. coli</i> Rosetta BL21 (DE3)	<i>F- ompT hsdSB(rB- mB) gal dcm (DE3) pRARE2</i> (<i>CamR</i>)	Protein expression	2.2.3

2.1.2. Cell Culture Media

Table 3. Cell culture media used in the following protocols.

Medium	Composition	Methods Section
Lysogeny Broth medium (LB)	1 % (w/v) Tryptone	2.2.1
	0.5 % (w/v) Yeast Extract	
	1 % (w/v) NaCl	2.2.2
Lysogeny Broth medium plates (LB agar)	1 % (w/v) Tryptone	2.2.1
	0.5 % (w/v) Yeast Extract	
	1 % (w/v) NaCl	2.2.2
	1.5 % (w/v) Agar	

2.1.3 Buffer Solutions

Table 4. Buffers used in the following protocols.

Buffer	Composition	Methods section
Buffer A	50 mM Hepes pH 7.5	2.2.3
	500 mM NaCl	
	25 mM Imidazole	
	10% (v/v) Glycerol	
Buffer B	50 mM Hepes pH 7.5	2.2.3
	500 mM NaCl	
	500 mM Imidazole	
	10% (v/v) Glycerol	
Desalting Buffer	50 mM Hepes pH 7.5	2.2.8
	150 mM NaCl	
	3% (v/v) Glycerol	
	1 mM DTT*	
Loading Buffer	200 mM Tris-HCl pH 6.8	2.2.5
	20% (v/v) Glycerol	
	10% (v/v) SDS	
Running Buffer	25 mM Tris-HCl pH 6.8	2.2.5
	192 mM Glycine	
	0.1% (v/v) SDS	

*added fresh in the purification day

2.2. Methods

2.2.1. Single-site mutagenesis

DszD_T62L and DszD_T62Y variants were obtained by single-site mutagenesis to replace Thr62 (T) with a leucine (L) and tyrosine (Y), respectively. Single-site mutagenesis reactions were performed in 20 μ L Polymerase Chain Reactions (PCR) using KOD HOT Start Polymerase® (Sigma-Aldrich) with 0.3 μ M of primers DszD_T62L_fw and DszD_T62L_rv and DszD_T62Y_fw and DszD_T62Y_rv (**Table 5**) and 10 ng of pET14_DszD (map in Annex, **Figure 36**), according to the program: 1- initial denaturing step 98 °C for 2'; 2 - 98 °C for 20'', 3 - 60 °C for 10'' and 4 - 72 °C for 3'; Repeat 18 cycles of step 2 to 4 for 18 times and a final extension step at 72 °C for 4'. One μ L of FastDigestion DpnI enzyme (ThermoFisher) was added to the PCR product and incubated for 60' at 37 °C prior to transformation in *Escherichia coli* DH5 α cells. The corrected codon substitution was confirmed by Sanger sequencing (Eurofins®).

Table 5. Primers used for single-site mutagenesis. Primers used for single-site mutagenesis to obtain DszD_T62L and DszD_T62Y variants. In purple are the mutation codons to substitute Thr62 to a leucine and a tyrosine.

Primer	Sequence
DszD_T62L_fw	GAAACCGAGCTGAAGGATAGCGCGGTAAAGCCCACCG
DszD_T62L_rv	CGGTGGGCTTTACCGCGCTATCCTTCAGCTCGGTTTC
DszD_T62Y_fw	GAAACCGAGCTGAAGGAATACGGTAAAGCCCACCG
DszD_T62Y_rv	CGGTGGGCTTTACCGCGTATTCCTTCAGCTCGGTTTC

2.2.2. Plasmid DNA Replication

E. coli DH5 α (**Table 2**) cells were transformed by adding 100 ng of plasmid and incubated on ice for 15'. The cells were heat shocked at 42 °C for 30" and cooled on ice for 5'. Cells were recovered in 600 μ L LB medium (**Table 3**) for 1h at 37 °C and 180 rpm (IKA[®] KS 4000 ic), and subsequently plated on LB agar plates (**Table 3**) supplemented with 50 μ g/mL kanamycin (KAN). The plates were incubated at 37 °C and 180 rpm (IKA[®] KS 4000 ic) overnight. The next day, one isolated colony was inoculated in 10 mL of LB medium supplemented with of KAN (50 μ g/mL). The pre-culture was incubated at 37 °C and 180 rpm (IKA[®] KS 4000 ic) overnight. The following day, the DNA was extracted and purified using the NZYMiniprep Kit (NZYTech), according to the manufacturer's protocol. The final plasmid concentration and its purity were assessed by measuring the absorbance at 260 nm and the A260nm/A280nm ratio, respectively, using a NanoDrop One (Thermo Fisher Scientific) spectrophotometer. The plasmids were stored at -20 °C.

2.2.3. Protein Expression

E. coli Rosetta BL21 (DE3) cells (**Table 2**) were transformed with pET14-DszD *wild type* (*wt*) or pET14-DszD_T62D, pET14-DszD_T62R, pET14-DszD_T62L, pET14-DszD_T62Y following the same protocol used for DNA replication in *E. coli* DH5 α . The next day, one isolated colony was inoculated in 100 mL of LB medium (**Table 3**) supplement with 50 μ g/mL KAN and 30 μ g/mL chloramphenicol (CAM), and incubated at 37 °C and 180 rpm (IKA[®] KS 4000 ic) overnight.

The following day, 20 mL of the pre-culture were added to 500 mL of LB medium supplemented with KAN (50 μ g/mL) and CAM (30 μ g/mL) in 2 L flasks. The flasks were incubated at 37 °C at 180 rpm (IKA[®] KS 4000 ic), until the optical density (OD) at 600 nm reached 0.8. At this point, 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to induce protein expression. The cells were incubated for 3 hours at 30 °C and 180 rpm (IKA[®] KS 4000 ic) and harvested by centrifugation at 4000 rpm for 20' at 4 °C (Beckman[®] Avanti J-26 XP, JLA- 81000 rotor). The supernatant was discarded, and the

pellet was resuspended in 3x the wet weight of the cells in volume of cold buffer A, flash frozen using liquid nitrogen and stored at -80 °C.

2.2.4. Protein Purification

Immobilized metal affinity chromatography (IMAC) was used to purify DszD *wt* and its variants. Both proteins were fused with an N-terminal 6x-His affinity tag and a thioredoxin solubility tag (Trx-tag). The cell suspensions were initially thawed at room temperature and then sonicated in a Sonoplus HD 2200/HD 3200 using a microtip MS73 (BANDELIN) for a total of 5' (30" ON; 30" OFF; 30% power). Afterwards, the lysate was clarified by centrifugation at 17000 rpm at 4 °C for 40' (Beckman ® Avanti J-26 XP, JA- 25.50 rotor). The supernatant was filtered through a low binding 0.45 µm PVDF Filter (CHROMAFIL® Xtra PVDF-20-25) before loading onto a Ni²⁺-charged HisTrap™ IMAC HP 5 mL column (Cytiva) using a peristaltic pump (Bio Rad® Econo Gradient). The column was previously equilibrated with 5 column volumes (CV) of Buffer A (**Table 4**). The clarified lysate was recirculated in the column for 40' at 2.5 mL/min before the flow-through (FT) was collected. For the washing and elution steps, the column was connected to an AKTAPrime Plus (GE Healthcare®) chromatography system and washed with 5 CV of ice-cold Buffer A. The protein was eluted in 5 mL fractions using different concentrations of imidazole (25 mL of 50 mM; 150 mM; 500 mM of imidazole) and a sample was taken in each purification step for SDS-PAGE gel analysis. After this analysis, the fractions corresponding to the elution with 125 mM imidazole and containing His_Trx_DszD were pooled for the next chromatographic step. The pooled fractions were diluted 3.5x in buffer without salt and imidazole to bring their final concentration to 150 mM and 35 mM, respectively, to ensure protein stability and Tobacco Etch Virus (TEV) protease activity. TEV protease was added to the solution in a 1:10 molar ratio relative to His_Trx_DszD and the mix was left overnight rotating at 4 °C.

The following day, the Ni²⁺-charged HisTrap™ IMAC HP 5 mL column (Cytiva) was equilibrated with 5 CV of Buffer A using a peristaltic pump (Bio Rad® Econo Gradient) and the protein was recirculated for 40' at 2.5 mL/min. DszD without His or Trx tags was collected in the FT. Then, the column was washed with 5 CV of buffer A and 5 CV of buffer B using a peristaltic pump (Bio Rad® Econo Gradient). Both His-Trx and His-TEV were eluted in buffer B. The FT was concentrated using a 10 kDa cut-off ultrafiltration device (Amicon ® Ultra-15). The protein concentration was measured on a NanoDrop One spectrophotometer (Thermo Fischer Scientific) using the theoretical molar extinction coefficient of DszD at 280 nm ($\epsilon_{280\text{nm}} = 16900 \text{ M}^{-1}.\text{cm}^{-1}$).

2.2.5. SDS-Page analysis

All expression and purification samples were analyzed on 12.5% (w/v) acrylamide resolving gels with 4% (w/v) of acrylamide stacking gels, since the molecular weight of all purified proteins was comprised between 20 and 80 kDa. To directly compare the level of protein expression, the same amount of cells were collected before (T0) and after the 3 hours of protein induction with IPTG (T3). The samples were centrifuged at 4000 rpm (Hitachi CT15RE centrifuge) for 2' at room temperature. The pellet was resuspended in 100 μ L of Buffer U (50 mM Tris pH 7.5, 100 mM NaCl, 6 M urea) and loading buffer (2% (w/v) SDS, 5% (v/v) glycerol, 50 mM Tris-HCl pH 6.8, 0.01% (w/v) bromophenol blue, 0.01% (w/v) 2-mercaptoethanol) was added to the samples. The gels were assembled into a Mini-PROTEAN® Tetra Cell System (Bio-Rad®) and ran at 200 V in running buffer (25 mM Tris-HCl pH 8.3, 192 mM Glycine, 0.1% (w/v) SDS) at room temperature, for approximately 45', stained with ProBlue Safe Stain (GiottoBiotech), and imaged in a Chemidoc Imaging System (Bio-Rad®).

2.2.6. Agarose gel electrophoresis of DNA analysis

DNA samples were analyzed on 1% (w/v) agarose gels containing GreenSafe (NZYTech) stain. Samples were prepared by adding DNA Loading Dye (0.5 M Ficoll 400, 100 mM EDTA; 1% (w/v) SDS; 3.6 mM bromophenol blue, 4.6 mM xylene cyanol). Gels were run at 90 V in Tris-Acetate-EDTA (TAE) buffer (40 mM Tris, 1mM EDTA, 0.11% (v/v) glacial acetic acid) for approximately 30'. The agarose gel was imaged using a Chemidoc Imaging System (Bio-Rad®).

2.2.7. Activity assays

DszD enzymatic activity was determined by measuring the decrease in absorbance at 340 nm due to the oxidation of NADH. The absorbance measurements were performed in 96 well plates (Falcon®) in a Synergy 2 Multi-Mode Plate Reader (BioTek) every 11" for 10' at 35 °C (**Figure 5**).

For activity measurements, the reaction contained 50 mM Hepes pH 7.5, 150 mM NaCl, 150 μ M FMN and 100 nM protein. Serial dilutions of NADH resulted in final concentrations ranging from 0.55 mM to 0.0808 mM. Absorbance values were converted to NADH concentrations by the Lamber-Beer law using the NADH extinction coefficient at 340 nm ($6220 \text{ M}^{-1} \cdot \text{cm}^{-1}$). The initial velocities were determined by linear regression of the initial linear phase of each activity curve. The kinetic parameters were calculated by plotting initial velocities versus the concentration of NADH.

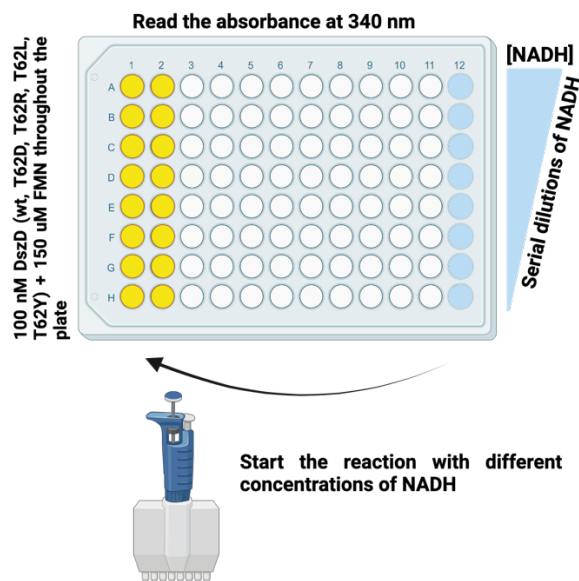


Figure 5. DszD's activity assay set-up. To the reaction mixture (100 nM DszD (wt, T62D, T62R, T62L, T62Y) and 150 μ M FMN (yellow) is added different concentrations of NADH (concentrations ranging from 0.55 mM to 0.081 mM) (blue), starting the enzymatic reaction. Its consumption is followed at a wavelength of 340 nm.

2.2.8. Size Exclusion Chromatography

A size exclusion chromatography column (Superdex75 10/300 GL, Cytiva) previously equilibrated with 50 mM Hepes pH 7.5, 150 mM NaCl, 3% (v/v) glycerol, 1 mM DTT using an AKTAPrimeTM (Cytiva) system was used to further purify DszD. The fractions containing DszD with the highest degree of purity were pooled and concentrated in a 10 kDa cut-off ultrafiltration device (Amicon® Ultra-0.5). The protein concentration was determined with a NanoDrop One spectrophotometer (Thermo Fischer Scientific) using the theoretical molar extinction coefficient at 280 nm of 16900 M⁻¹.cm⁻¹.

2.2.9. Crystallization Screenings

Initial crystallization conditions were screened using the following commercially available crystallization kits: Morpheus[®] MD1-46 (Molecular Dimensions[®]), JCSG+ 1-2 (Hampton Research[®]) and Crystal Screen Cryo (Hampton Research[®]). The crystallization screenings were performed using an Oryx4 robot (Douglas Instruments) in 96-wells crystallization plates (3 Well Low Profile Crystallization Plate (Swissci)). DszD *wt* was crystallized using the sitting-drop vapour diffusion method at a concentration of 300 μ M in the absence and in the presence of FMN at a concentration of 600 μ M. The protein/precipitant ratio was 1:1 in drops of one μ L. The crystallization plates were incubated at 20 °C and were checked after 24 hours, two days, three days of incubation, one week after incubation and then every week during the month.

The crystallization conditions that had crystals were scaled up to sitting-drop vapor diffusion 24-wells crystallization plates (Cryschem plates, HamptonResearch ®). The protein/precipitant ratio was 1:1 in drops of 2 μL with 300 μL reservoirs. The crystallization plates were incubated at 20 °C and were checked after 24 hours, two days, three days of incubation, one week after incubation and then every week during the month.

RESULTS

3.1. Expression and purification of DszD *wild type* and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y)

Both DszD *wt* and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y) were cloned into a pET14 plasmid under the control of the IPTG-inducible T7 promotor (map in Annex, **Figure 36**) and were expressed in *E. coli* Rosetta (DE3) cells. To verify the level of protein expression, samples were collected at different time points, before IPTG induction (T0) and 3 hours after IPTG induction (T3) and analyzed by SDS-PAGE. The protein expression of both His-Trx-DszD *wt* and variants was confirmed by the appearance of a band with the expected molecular weight of 34 kDa (Trx molecular weight is 14 kDa and DszD has 20 kDa) after IPTG induction (**Figure 6**).

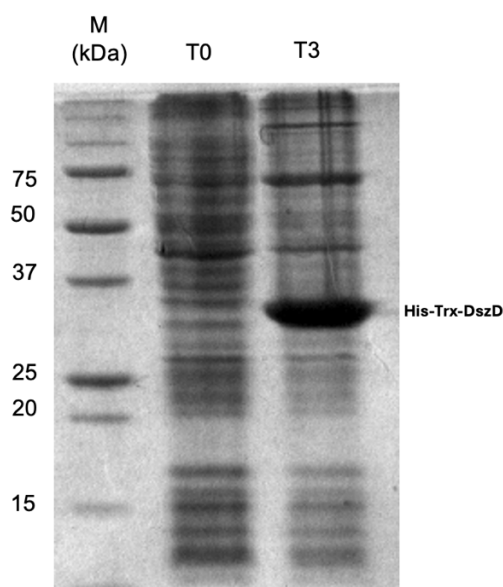


Figure 6. DszD *wt* protein expression. SDS-PAGE analysis of the fractions collected before (T0) and 3 hours after IPTG induction (T3). Expression of DszD *wt* was confirmed by the appearance of a band with the expected molecular weight of 34 kDa after IPTG induction at the T3 lane. The first lane (**M (kDa)**) corresponds to the Precision Plus Protein™ Standards molecular mass marker (Bio-Rad®).

After protein expression, DszD *wt* and variants were purified in two sequential IMAC steps. Both DszD *wt* and variants were fused to a thioredoxin solubility tag (Trx-tag) and a histidine affinity tag (6xHis-tag). The Trx-tag was included to enhance solubility, while the 6xHis-tag facilitated purification via IMAC using Ni^{2+} ions. A TEV protease recognition sequence (ENLYFQ▼S) was inserted between the affinity/solubility tags and the DszD sequence. This allowed for the removal of the tags through cleavage with TEV protease,

resulting in tag-free DszD, which is preferred to perform biochemical assays and crystallization experiments.

The DszD purification protocol had been previously optimized in the laboratory. Briefly, the first purification step is required to isolate the protein of interest from other soluble cellular components and His-Trx-DszD protein was eluted using a stepwise increase in the imidazole concentration: 50 mM, 125 mM and 500 mM imidazole (**Figure 7a**). The eluted fractions were analyzed by SDS-PAGE, and His-Trx-DszD protein was predominantly eluted with 125 mM imidazole (**Figure 7b**). The fractions containing His-Trx-DszD were pooled and diluted to 150 mM NaCl with buffer with no salt and imidazole prior to overnight incubation with TEV protease. This dilution is necessary for optimal protease activity, which requires a salt concentration below 200 mM. In addition, the protease was added at a 1:10 molar ratio relative to the target protein, whose concentration was determined from its absorbance at 280nm. After overnight incubation, the pooled fractions were loaded into the IMAC column and DszD was collected in the flow-through, while the affinity and solubility tags, as well as the TEV protease and uncleaved protein (His-Trx-DszD) bound to the column and were eluted with 500 mM imidazole. SDS-PAGE analysis confirmed the purity of DszD eluted in the flow through with the predicted molecular weight of 20 kDa (**Figure 7c**). Additionally, a band of approximately 40 kDa was observed in the gels of all purified samples, likely representing a dimer (**Figures 7c and 8**). Furthermore, the intensity of this band decreases with increasing concentrations of reducing agent.

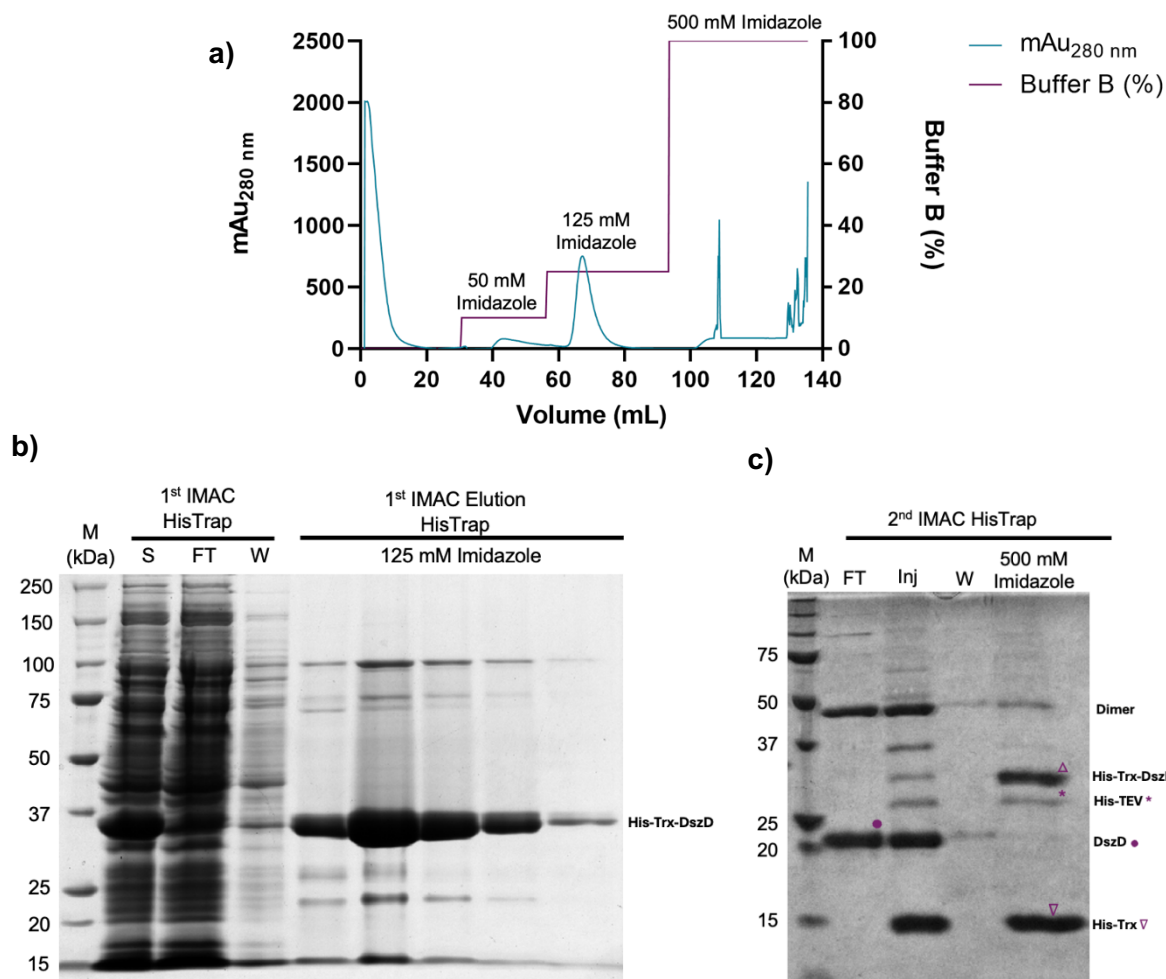


Figure 7. DszD wt purification by Immobilized Metal Affinity Chromatography. a) Chromatogram of the 1st IMAC purification step of Trx-His-DszD wt from the soluble extract of *E. coli* Rosetta (DE3). Trx-His-DszD wt was eluted in 125 mM imidazole (25% Buffer B). In blue is represented the absorbance (mAu) at 280 nm and in purple is the percentage of Buffer B used in the elution step (10%; 25% and 100%). b) SDS-PAGE analysis of the 1st IMAC purification step of Trx-His-DszD wt. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad); **S**, filtered supernatant obtained after cell lysis; **FT**, Flow-through of the loading step of the supernatant onto a Ni²⁺ charged HisTrap™ IMAC HP (Cytiva) 5 mL column; **W**, wash of the column with buffer A (**Table 4**); **1st IMAC elution HisTrap**, Trx-His-DszD wt elution fractions with 125 mM imidazole. c) SDS-PAGE analysis of the 2nd IMAC purification step of DszD wt. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad), **FT**, Flow-through obtained after recirculating the DszD wt cleavage mixture on a Ni²⁺ charged HisTrap column (Cytiva); **Inj**, DszD wt cleavage mixture after overnight incubation; **W**, wash of the column with buffer A (**Table 4**); **500 mM imidazole**, 500 mM imidazole elution step to elute the bound proteins: His-Trx, His-TEV and His-Trx-DszD.

Both DszD wt and its variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y) were successfully purified using the two IMAC step approach (**Figure 8**).

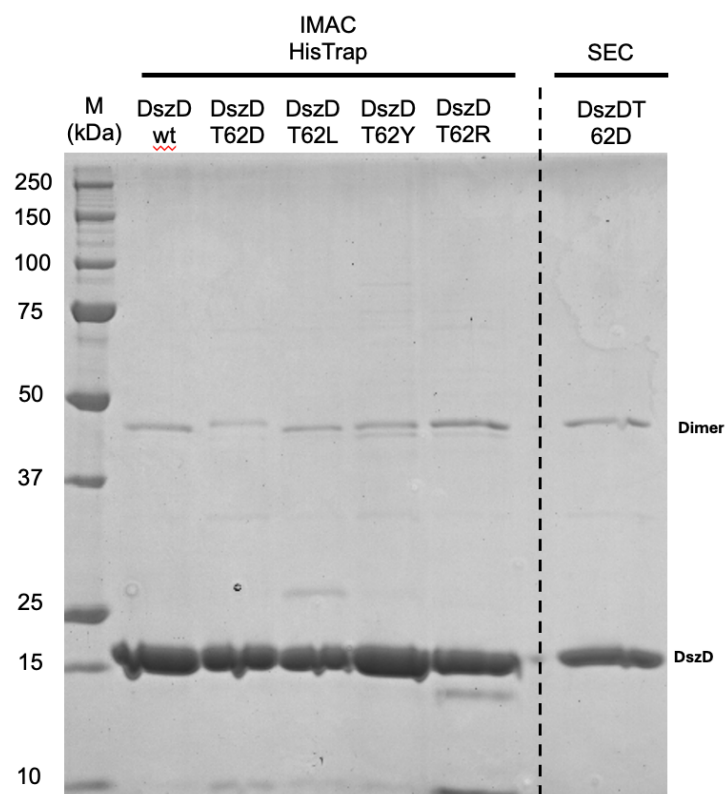


Figure 8. SDS-PAGE gel of DszD wt and variants purified using two IMAC purification steps. SDS-PAGE analysis of the purified DszD wt and variants - DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y - after the two steps of purification using IMAC and of DszD_T62D after SEC purification. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad).

3.2. Activity assays of DszD wt and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y)

To assess whether the amino-acid substitutions (T62D, T62R, T62L, and T62Y) led to an increase on the enzymatic activity of DszD, we conducted enzymatic assays. In this assay, we monitored the oxidation of NADH to NAD⁺, coupled with the concomitant reduction of FMN to FMNH₂ catalysed by DszD. For this assay, we took advantage of the differential light absorption propriety of NADH and NAD⁺ at 340 nm, since NADH strongly absorbs light at 340 nm and its oxidated form does not. This spectral difference allows the quantitative measurement of NADH consumption at 340 nm by monitoring the decrease in absorbance at 340 nm over time.

To optimize the enzymatic assay conditions, we first performed a preliminary assay with DszD wt to determine the linearity of the enzymatic reaction with increasing NADH concentrations, as well as to determine the appropriate protein concentration for the activity assay. For this, we tested different protein concentrations (3 nM, 12.5 nM, 50 nM,

100 nM and 200 nM), that were incubated with 150 μ M FMN before initiating the reaction with either 0.2 mM or 0.4 mM NADH. The reaction was monitored by measuring the absorbance at 340 nm. The initial velocity for each enzyme concentration was extrapolated from the linear portion of the progress curves. The results of this assay demonstrate that the reaction rate increases linearly with DszD concentrations up to 100 nM for the different NADH concentrations tested. At 200 nM DszD, a deviation of linearity was observed, likely due substrate depletion or product inhibition (**Figure 9**).

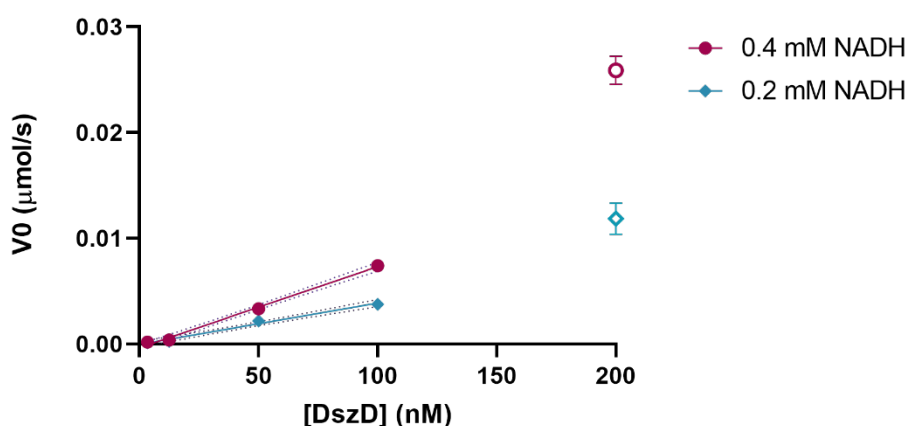


Figure 9. Linear dependence of the initial velocity with DszD *wt* enzyme concentration, at different NADH concentrations. The reaction rate is linear with DszD concentration up to 100 nM for both NADH concentrations (0.2 mM and 0.4 mM). The reaction was monitored by the absorbance at 340 nm. Error bars represent the standards deviation of technical duplicates measurements.

To establish the NADH concentration range for the activity assays, it was essential to know the linear range of the absorbance at 340 nm in the Synergy 2 Multi-Mode Plate Reader (BioTek) where our measurements were conducted. According to a study performed by BioTek (Held, Paul. (2021, September 1)) the absorbance increases linearly with NADH concentrations up to 0.6 mM. Concentrations of NADH above 0.6 mM result in signal saturation of the spectrometer. To ensure that all the measurements remain within the linear range of the detection of the instrument, the concentrations of NADH used for the assay were kept below the 0.6 mM limit (**Figure 10**).

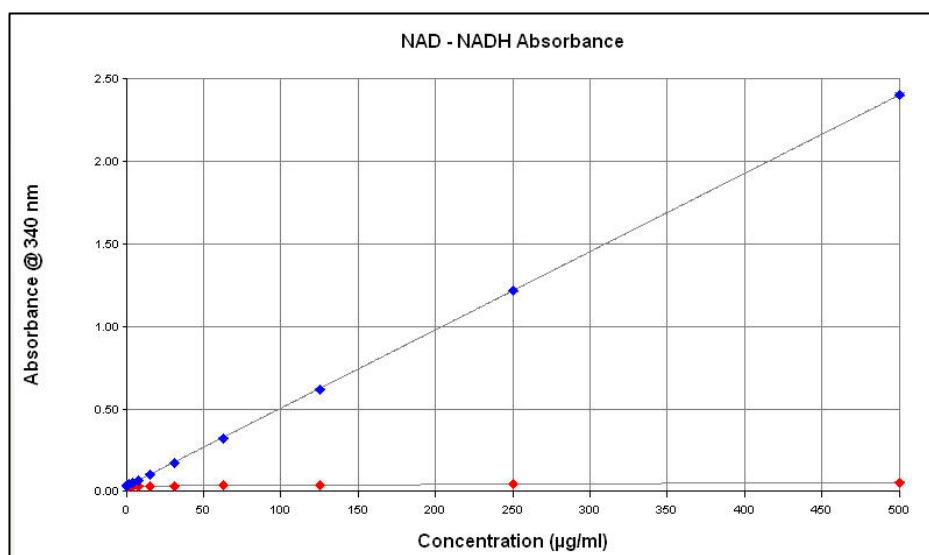


Figure 10. NADH and NAD⁺ concentration curve measured using the absorbance at 340 nm. The absorbance of NADH (blue dots) and NAD⁺ (red dots) at different concentrations was measured at 340 nm using the Synergy 2 Multi-Mode Plate Reader (BioTek). Data was taken from an application note made available by BioTek ("Determination of NADH Concentrations with the Synergy™ 2 Multi-Detection Microplate Reader using Fluorescence or Absorbance," BioTek Application Note, 2006).

Activity assays were performed to determine if the DszD variants (T62D, T62R, T62L, and T62Y) were catalytically more active than the *wt* protein. The enzymatic reaction was initiated by adding NADH at different concentrations, ranging from 0.08 mM to 0.5 mM, to a pre-incubated mixture of 100 nM protein and 150 µM FMN. To remain below the instrument's upper limit, the maximum concentration of NADH used was 0.5 mM. The reaction was monitored by measuring the absorbance at 340 nm, indicative of NADH consumption by the enzyme (**Figure 11**). The enzymatic assay was performed for all the variants using the same conditions as for the *wt* protein.

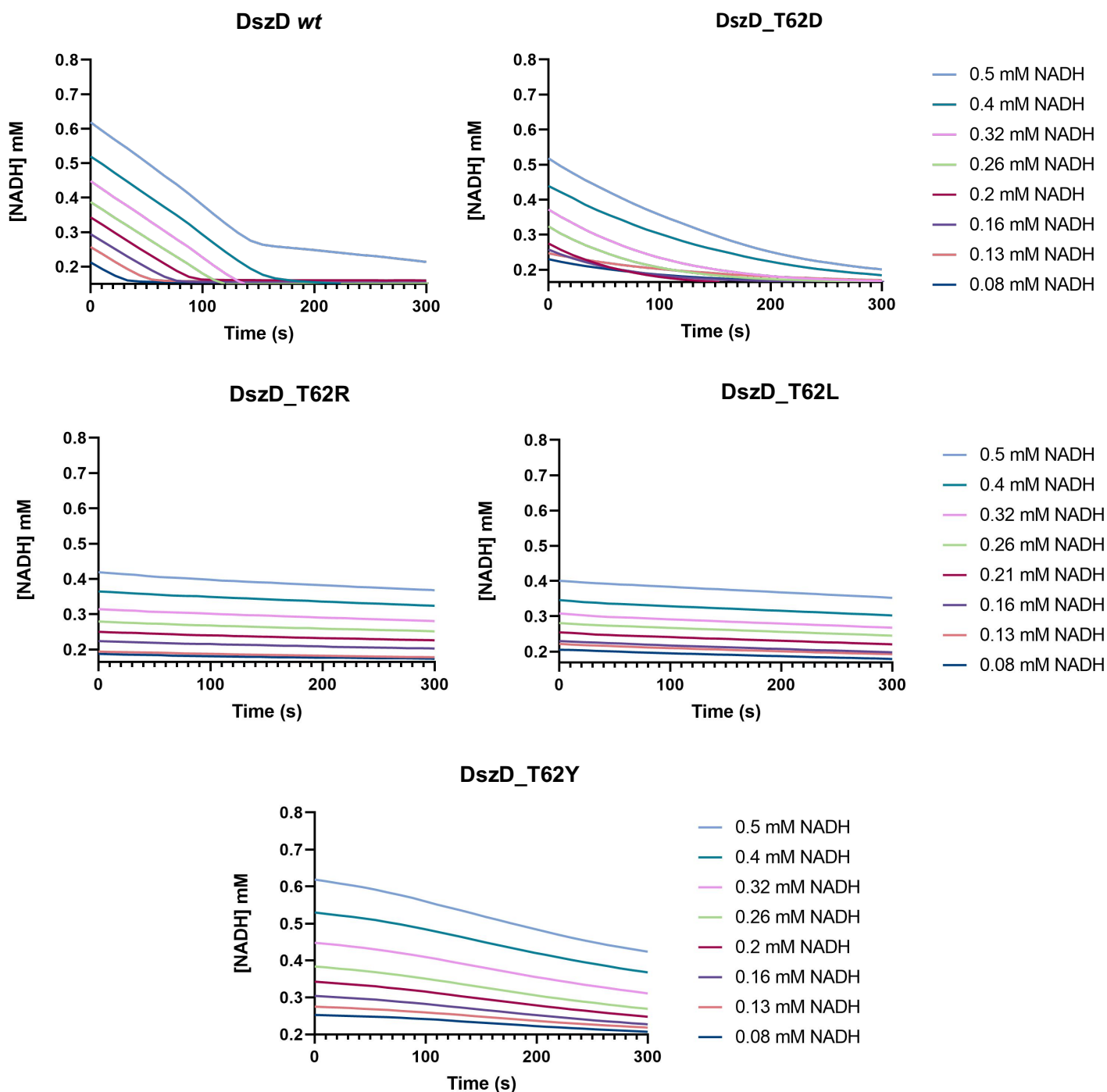


Figure 11. Consumption of NADH by DszD wt and variants DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y as a function of time (300 s). The absorbance was measured at 340 nm and then converted to NADH concentrations by the Lambert-Beer law ($\epsilon_{340 \text{ nm}} = 6220 \text{ M}^{-1} \cdot \text{cm}^{-1}$). Each curve represents a reaction started with different NADH concentrations ranging from 0.08 mM to 0.5 mM.

Nearly all enzymes follow a Michaelis-Menten kinetics model and to determine the kinetic parameters that characterize enzyme activity (k_{cat} and K_m), we kept the concentration of the enzyme constant to allow for the maximum velocity of the reaction (V_{max}) to be reached and also assumed that the enzymatic activity of DszD occurs under steady-state conditions, where the concentration of the enzyme/substrate (ES) complex is constant

because the formation of ES equals its dissociation. K_m and k_{cat} are parameters that provide insight into the affinity and catalytic efficiency of enzymes, respectively. To determine these kinetic parameters to both DszD *wt* and variants, the initial velocity (V_0) of each NADH concentration curve was extrapolated from the linear portion of the progress curves (**Figure 12**). We then represented V_0 vs NADH concentration and applied the equation $V_0 = (E_t \times k_{cat} \times [S]) / (K_m + [S])$, where E_t is the enzyme concentration, $[S]$ the substrate concentration, and V_0 the initial velocity, to calculate both K_m and k_{cat} . After obtaining these kinetics parameters, we were able to compare the activity of the different variants. **Table 6** summarizes the k_{cat} , K_m , and maximum velocity (V_{max}) values obtained for each variant.

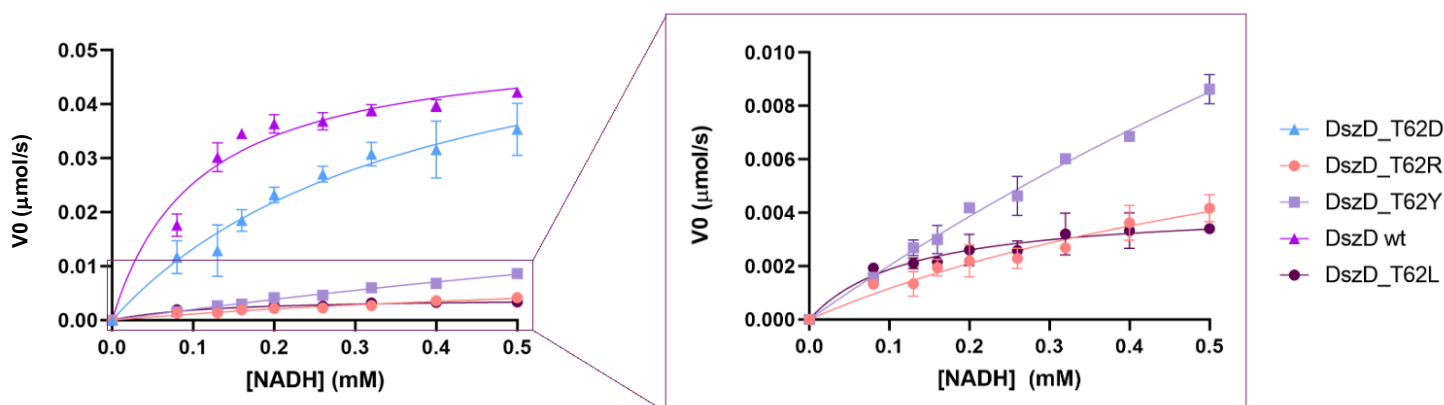


Figure 12. The impact of substitutions of residue 62 in the enzymatic activity of DszD. V_0 was extrapolated using the linear portion of the curve for each concentration of NADH (0.08 mM to 0.5 mM). On the right side the initial part of the curves was amplified for more detail. From these curves the kinetics parameters (K_m , V_{max} and k_{cat}) for both DszD *wt* and variants were determined using the Prism 8 (GraphPad Software). Error bars represent the standards deviation of technical duplicates measurements.

The K_m parameter, which corresponds to the substrate concentration at which the reaction rate is half of the maximum velocity (V_{max}), varied significantly among the variant. Both DszD *wt* and DszD_T62L displayed very similar K_m values (0.11 mM and 0.12 mM, respectively) that were also the lowest of the series. This indicates that these two proteins have the highest substrate affinity, and that replacing Thr62 with a leucine had a minimal effect on substrate binding. In contrast, DszD_T62Y exhibited the highest K_m (2.07 mM) among all the variants, indicating a significant decrease in substrate affinity.

Regarding the k_{cat} , which describes the number of substrate molecules that are converted to product by an enzyme per unit of time when the enzyme is fully saturated with substrate, it ranged from 42.38 s⁻¹ to 635.60 s⁻¹. In this case, DszD_T62D displayed

the highest k_{cat} (635.60 s^{-1}), surpassing the *wt* protein (520.50 s^{-1}), suggesting that this amino acid substitution increases the catalytic rate of the enzyme. In contrast, DszD_T62L displayed the lowest k_{cat} (42.38 s^{-1}), indicating a significant decrease in the catalytic rate.

Overall DszD *wt* exhibited the highest catalytic efficiency ($4933.59 \text{ mM}^{-1} \cdot \text{s}^{-1}$), which is described by k_{cat}/K_m , being followed by DszD_T62D ($1670.87 \text{ mM}^{-1} \cdot \text{s}^{-1}$). Despite having the highest catalytic rate, the high K_m of DszD_T62D resulted in a lower catalytic efficiency compared to the *wt* protein. The variants with lowest catalytic efficiency were DszD_T62R and DszD_T62Y due to their high K_m values.

Table 6. Kinetics parameters (k_{cat} , K_m and V_{max}) of DszD *wt* and variants, DszD_T62D, DszD_T62R, DszD_T62L and DszD_T62Y. The 95% profile likelihood and goodness of fit are also indicated.

	DszD <i>wt</i>	DszD_T62D	DszD_T62R	DszD_T62L	DszD_T62Y
Best fit values					
Et	0.00010	0.00010	0.00010	0.00010	0.00010
$k_{\text{cat}} (\text{s}^{-1})$	520.50	635.60	106.00	42.38	437.60
$K_m (\text{mM})$	0.11	0.38	0.81	0.13	2.07
$V_{\text{max}} (\mu\text{M} \cdot \text{s}^{-1})$	0.052	0.064	0.011	0.0042	0.044
95% CI (profile likelihood)					
k_{cat}	469.40 to 583.00	483.70 to 931.20	61.29 to 452.50	34.35 to 54.93	260.4 to 1474.00
K_m	0.074 to 0.15	0.22 to 0.71	0.34 to 4.64	0.060 to 0.24	1.09 to 7.84
Goodness of fit					
R squared	0.9641	0.9389	0.9125	0.8888	0.9844

3.3. Effect of FMN on DszD oligomerization

DszD provides FMNH₂ to two different monooxygenases, DszA and DszC. Changes in the oligomeric state of DszD could regulate the transfer of the reduced flavin to each monooxygenase during the desulfurization process and influence the overall enzymatic activity. Therefore, we assessed the impact of FMN in the oligomerization state of DszD by incubating 50 μM (1 mg/mL) DszD *wt* with increasing concentrations of FMN (25 μM , 50 μM , 100 μM and 200 μM) prior to separation on an analytical SEC column (Superdex200 Increase 3.2/300 (Cytiva)) previously equilibrated with 50 mM Hepes pH 7.5, 150 mM NaCl, 3% (v/v) glycerol, 1 mM DTT (**Table 4**), allowing the analysis of both the oligomeric state(s) of DszD and the possible impact of FMN on oligomerization.

To accurately predict the molecular weight of the different oligomeric states of DszD, both alone and in the presence of FMN, a calibration curve with protein standards was performed, including bovine serum albumin (BSA), ovalbumin, chymotrypsin, and RNase on the same SEC column (Superdex200 Increase 3.2/300 (Cytiva)) using the same conditions. By correlating the distribution coefficient (K_{av}) with the known molecular

weights of the standard proteins, it is possible to estimate the molecular weight (MW) of the different oligomeric forms of DszD. The elution volumes of the protein standards were 1.51 mL for aldolase (MW = 158 kDa), 1.66 mL for BSA (MW = 66.43 kDa), 1.78 mL for ovalbumin (MW = 44.5 kDa), 2.00 mL for chymotrypsin (MW = 25 kDa), and 2.08 mL for RNase (MW = 13.7 kDa) (**Figure 13a**). The K_{av} and the predicted molecular weight of each oligomeric form of DszD *wt* in the absence and in the presence of increasing FMN concentrations (25 μ M, 50 μ M, 100 μ M and 200 μ M) are represented in **Table 7**.

The elution volume of DszD indicates that in the absence of FMN, the apparent molecular weight of DszD *wt* is approximately 108 kDa, which would be consistent with a pentameric arrangement. With the increase of [FMN], there is a clear shift of this main elution peak towards higher elution volumes, as well as a formation of a shoulder in the left region of the elution profile of DszD (**Figure 13c**). These observations probably indicate a change in the oligomeric state and/or the overall folding of DszD as the FMN concentration increases. Given that these findings are conflicting with the dimer predicted by the 3D structure model (Sousa *et al.* 2016), we used AlphaFold to predict the assembly of DszD as a pentamer and although the algorithm could produce a model, the confidence of the prevision was very low (ipTM = 0.27 compared to ipTM=0.78 for the dimer). Hence, it seems unlikely that DszD assembles as pentamer.

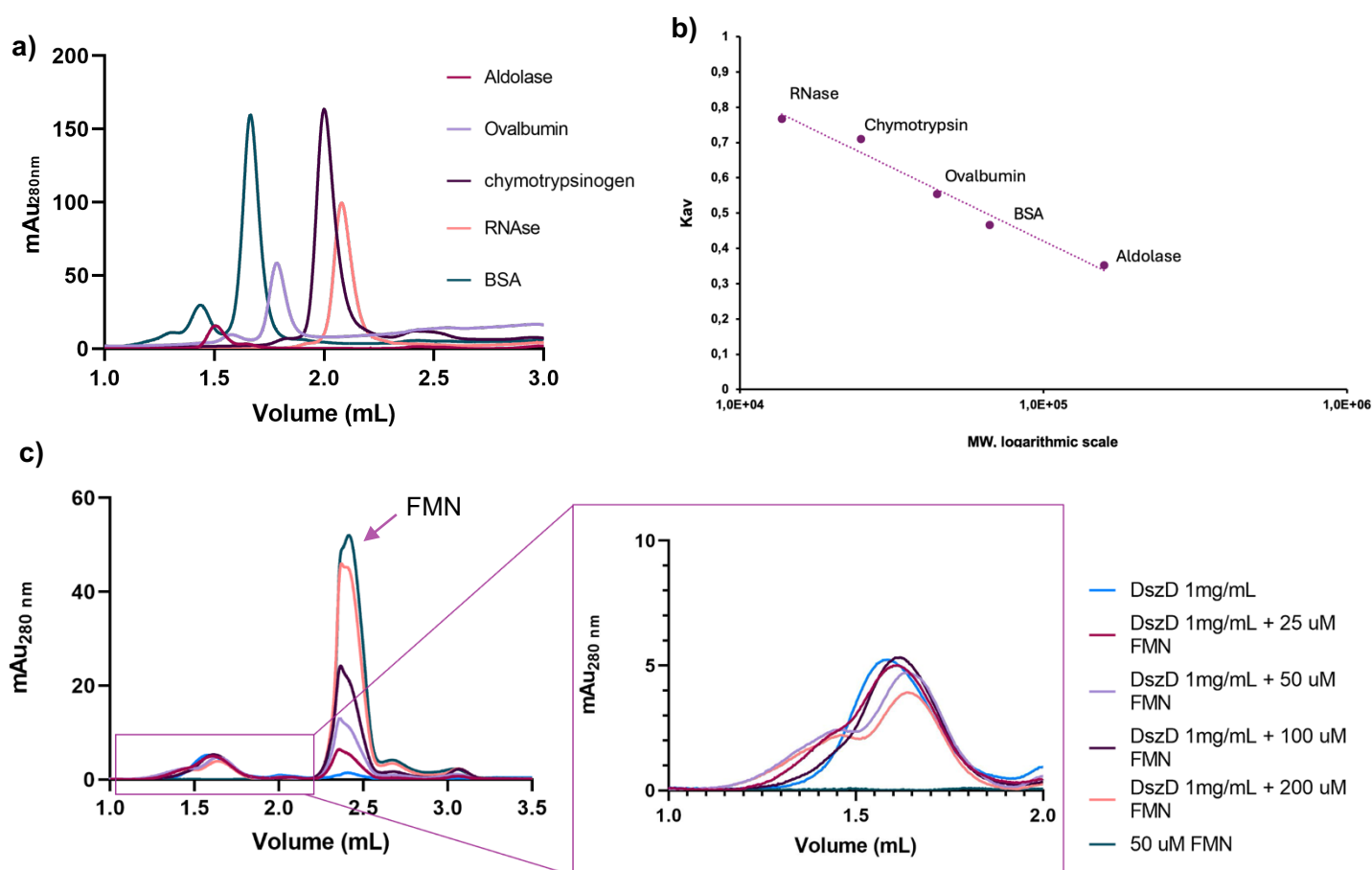


Figure 13. The effect of FMN on the oligomerization of DszD wt. a) Elution profiles of the protein standards that includes aldolase (158 kDa), BSA (66.43 kDa), ovalbumin (44.5 kDa), chymotrypsin (25 kDa) and RNase (13.7 kDa). b) Calibration curve shows distribution coefficient (K_{av}) vs log molecular weight. The equation of the calibration curve obtained is $y = -0,181\ln(x) + 2,5052$ with $R^2 = 0,9758$. K_{av} was calculated using the following equation: $K_{av} = (V_e - V_0) / (V_c - V_0)$, where V_0 is the column volume void and V_c is the total column volume. c) Elution profile of DszD wt in the absence and in the presence of increasing concentrations of FMN (25 μ M, 50 μ M, 100 μ M and 200 μ M). The region between 1 and 2 mL was amplified for more detail on the right side.

Table 7. K_{av} and the predicted molecular weight of DszD wt. K_{av} and the corresponding predicted molecular weight of DszD wt in the absence and presence of increasing FMN concentrations (25 μ M, 50 μ M, 100 μ M and 200 μ M).

[FMN] (μ M)	K_{av}	MW DszD (kDa)
0	- 0.41	- 108.90
25	0.31 0.43	183.23 96.58
50	0.31 0.44	183.23 89.15
100	0.32 0.47	176.04 75.96
200	0.32 0.45	176.04 85.65

3.4. DszD's AlphaFold structure prediction

Given the conflicting estimation of the molecular weight of DszD wt by SEC (**Figure 13c** and **Table 7**) with the available structural model (40kDa, dimer), we decided to re-analyze the structure of DszD. The initial model by Sousa *et al.* (2016) did not contain the first 29 amino acids of the DszD sequence because it was predicted using the crystal structures of PheA2, Hpac_{St} and HpaC_{Tt} as templates. Although the catalytic pockets of these proteins are similar, they lack the N-terminal stretch present in DszD and, for this reason, the predicted 3D structure of DszD does not feature the first 29 residues. We used AlphaFold, that was not available when the model was obtained, to predict the 3D structure of full-length DszD and it suggested that the first 27 amino acids of DszD are probably disordered since the algorithm is unable to predict their 3D structure with confidence (pLDDT < 50, very low confidence) (**Figure 14**). Hence, the presence of a disordered N-terminal may be the reason for the abnormal elution volume of DszD wt. To test this hypothesis, we generated a truncated version of the protein (Δ N-DszD). Independent *in vivo* assembly (IVA) was used to delete the first 81 nucleotides from the *dszd* gene coding for the 27 N-terminal amino acids (**Table 8**) and successful deletion was confirmed by Sanger sequencing. Next, we expressed and purified Δ N-DszD and analyzed it by SEC (**Figure 15**). The elution volume of Δ N-DszD is 2.00 ml (K_{av} of 0.71) (vs 1.58 ml and K_{av} 0.41 for the full-length protein) which corresponds to an apparent

molecular weight of approximately 20kDa (**Table 9**) suggesting that in the absence of the N-terminal DszD is a monomer in solution (18kDa). Taken together, our SEC results indicate that DszD was eluted with higher apparent molecular weight than expected from its amino acid sequence because its N-terminal stretch is disordered, as predicted by AlphaFold.

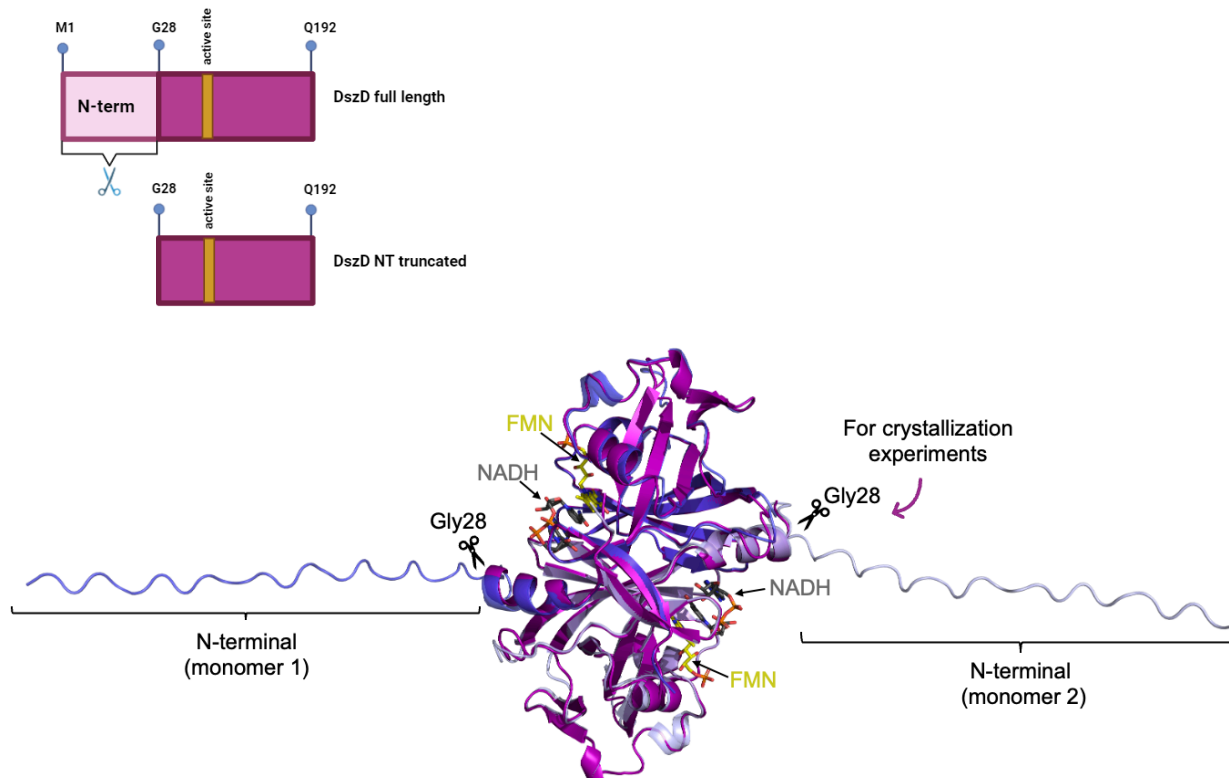


Figure 14. DszD structure predicted by AlphaFold merged with Sousa et al. (2016) model. The full-length structure predicted by AlphaFold shows a possibly disordered N-terminal. For future crystallization experiments, a truncated version was generated by removing the first 27 residues. In the AlphaFold structure, monomer 1 is in dark blue and monomer 2 is in light blue. The structure from Sousa *et al.* (2016) is represented in purple. The DszD structures are in cartoon representation. The FMN (yellow) and NADH (grey) molecules are represented as sticks.

Table 8. Primers used in IVA cloning. Primers used to delete the N-terminal region using IVA cloning. The homologous region of the primer is shown in purple.

Primer	Sequence
ΔN-DszD_fw	CACCGGCGCAGATACCTTTGCCCATGGCGCCC
ΔN-DszD_rv	GGTATCTGCGCCGGTGATTAC

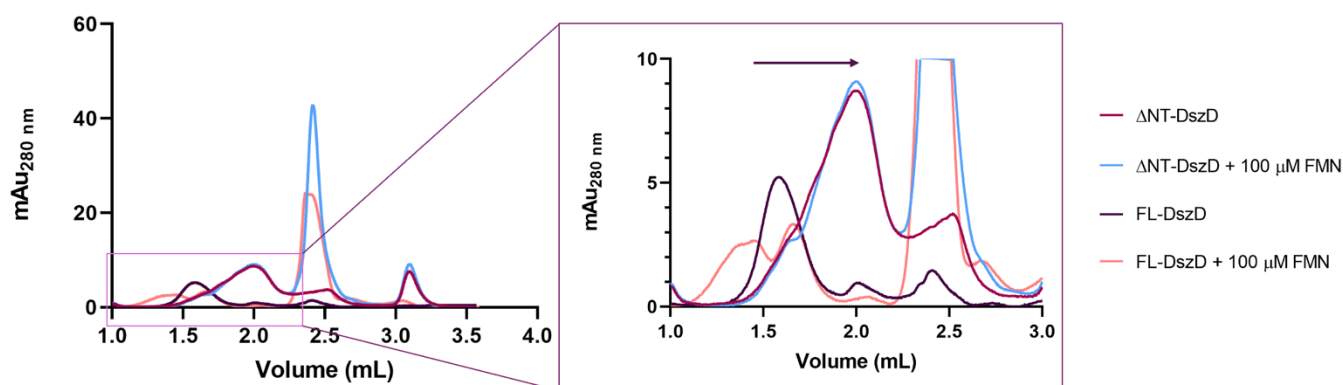


Figure 15. SEC elution profiles of Δ NT-DszD and FL-DszD. Analytical SEC chromatograms of Δ NT-DszD and FL-DszD in the absence and in the presence of 100 μ M FMN. The region between 1 and 3 mL is magnified for more detail on the right inset.

Table 9. K_{av} and the predicted molecular weight of Δ NT-DszD. The elution volume and the corresponding Δ NT-DszD molecular weight in the absence and presence of 100 μ M FMN.

	K_{av}	MW DszD (kDa)
Δ nt-DszD	0.71	20.21
Δ nt-DszD + 100 μ M FMN	0.72	19.42

3.5. Crystallization of DszD *wt*

There is currently no three-dimensional structure of DszD from *R. erythropolis* IGTS8, either in its apo-form or complexed with NADH and/or FMN. Therefore, in order gain deeper insights into the overall protein structure, the enzyme's active site and substrate binding pocket at the atomic level, we decided to crystallize DszD *wt*. Further, determining the structure of DszD could provide valuable information on the catalytic mechanism of both the *wt* protein and the variants and better guide substitutions to enhance its catalytic power.

To identify crystallization conditions, we conducted crystallization screens using commercially available crystallization kits, Morpheus[®] MD1-46 (Molecular Dimensions) and JCSG+ 1-2 (Hampton Research). The crystallization screens were set up using DszD *wt* purified by IMAC, at a concentration of 6 mg/mL (300 μ M), in the presence of a 2-fold molar concentration of FMN (600 μ M). The protein/precipitant volume ratio was 1:1 in drops of 1 μ L. Several conditions in the Morpheus[®] MD1-46 crystallization kit yielded crystals, but we focused on two promising crystallization conditions that produced needle-like crystals. The crystallization conditions were: C3 (0.09 M NPS, 0.1 M buffer

system 1 pH 6.5, 30% (v/v) precipitant mix 3) and G3 (0.12 M carboxylic acids, 0.1 M buffer system 1 pH 6.5, 30% (v/v) precipitant mix 3) (**Figure 16**) (**Table 13** in the annex). However, these crystals appeared to be too small for X-ray data collection, and to obtain larger crystals these conditions were reproduced and scaled-up in 24-well crystallization plates (Cryschem plates, Hampton Research) using 1 μ L of purified protein and 1 μ L of reservoir solution per drop. The crystals appeared after one day and grew over the following days. However, during data collection at the ALBA Synchrotron (Cerdanyola del Vallès, Spain), we observed that the crystals were not protein crystals as initially assumed, but rather FMN salt crystals.

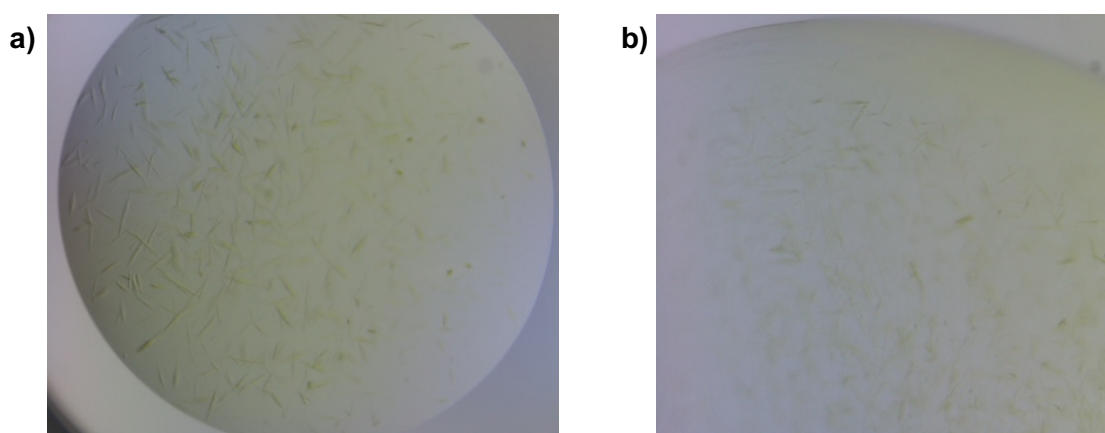


Figure 16. Crystals obtained from two Morpheus® MD1-46 (Molecular Dimensions) conditions. Crystallization condition: 6 mg/mL DszD *wt* with 600 μ M FMN in vapor diffusion in 24-well crystallization plates (Cryschem plates, Hampton Research). a) condition C3 (0.09 M NPS, 0.1 M buffer system 1 pH 6.5, 30% (v/v) precipitant mix 3); b) condition G3 (0.12 M carboxylic acids, 0.1 M buffer system 1 pH 6.5, 30% (v/v) precipitant mix 3) (**Table 13** in the annexes).

For quality control and for removing excess of FMN, DszD *wt* at a concentration of 6 mg/mL, in the presence of FMN at twice the protein concentration (600 μ M FMN) was injected into a SEC column (Superdex75 10/300 GL, Cytiva). This ensured that the sample used for the subsequent crystallization screenings was as homogeneous as possible (**Figure 17**). After SEC, new crystallization screenings with DszD *wt* were set up using Morpheus® MD1-46 (Molecular Dimensions), JCSG+ 1-2 (Hampton Research) and Crystal Screen Cryo (Hampton Research). Various conditions of Morpheus® MD1-46 (Molecular Dimensions) and Crystal Screen Cryo (Hampton Research) yielded crystals, with the latter having the most promising looking crystals (**Figure 18**). The crystallization condition from Crystal Screen Cryo was: B11 (0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% (v/v) Polyethylene glycol 400). To obtain more and larger crystals from the B11 crystallization condition of Crystal Screen Cryo for subsequent X-ray data collection, we attempted to scale up the crystallization condition

in 24-well sitting drop crystallization plates (Cryschem plates, Hampton Research), but we didn't obtain any crystals. Hence, the crystals from the original 96-well crystallization plates (3 Well Low Profile Crystallization Plate (Swissci)) were harvested for subsequent data collection. Once again, we verified that the crystals were not protein, but salt crystals.

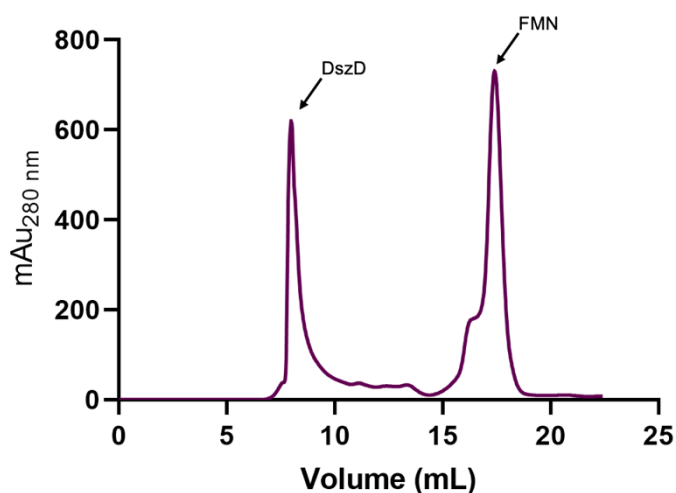


Figure 17. SEC chromatogram of DszD *wt* in the presence of FMN. DszD *wt* at a concentration of 6 mg/mL was incubated with 600 μ M FMN and injected into a SEC column (Superdex75 10/300 GL, Cytiva). The purple line represents the absorbance (mAu) at 280 nm.

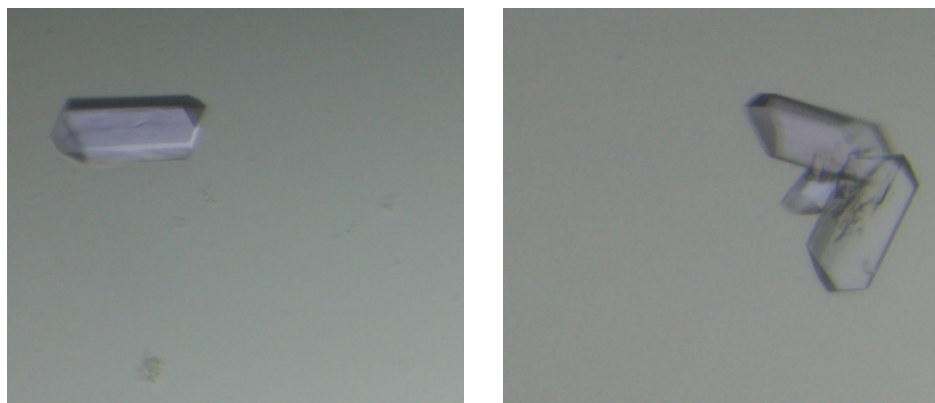


Figure 18. Crystals obtained from B11 condition of Crystal Screen Cryo screen (Hampton Research). Crystallization condition: 6 mg/mL DszD in vapour diffusion with reservoir solution: 0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% (v/v) PEG 400 in 96-well crystallization plates (3 Well Low Profile Crystallization Plate (Swissci)).

DISCUSSION

DszD is a flavin reductase from the 4S pathway that catalyses the reduction of FMN to FMNH₂ through the oxidation of NADH to NAD⁺. This enzyme is responsible for supplying FMNH₂ two other enzymes from this pathway, DszC and DszA. Given the key function of this enzyme in the 4S pathway, one strategy proposed to increase the reaction rate at which the bacterial cells remove sulfur in BDS and make this process viable for industrial application is to increase the catalytic power of DszD through rational mutagenesis of important residues for its activity. Previous studies performed by Kamali *et al.* (2010) identified Thr62 as a key residue in the rate-limiting step of the reduction of FMN to FMNH₂. To accelerate this reaction, Sousa *et al.* (2016) and Ferreira *et al.* (2017) performed *in silico* studies by systematically replacing the Thr62 key residue. Out of the 18 different amino acids substitutions studied, we selected three substitutions (T62D, T62L and T62Y) for their potential to enhance DszD catalytic power and one substitution for its predicted negative effect (T62R).

The kinetic analysis performed on both DszD *wt* and variants (DszD_T62D, DszD_T62R, DszD_T62L, DszD_T62Y) provided significant insights into the role of residue 62 in both substrate binding and catalytic activity. Amongst all variants, the T62L substitution showed a similar K_m to DszD *wt*, K_m= 0.1248 mM and 0.1054 mM, respectively, which suggests similar substrate affinity. This could be due to the similar size and spatial orientations of threonine and leucine in the active site, which according to Ferreira *et al.* (2017), could allow NADH to be “locked” in a favorable manner (**Figure 12** and **Table 6**).

Notably, DszD_T62D exhibited the highest catalytic rate (k_{cat} of 635.6 s⁻¹) among all variants, including the *wt* protein (**Figure 12** and **Table 6**). The negatively charged aspartate at position 62 may facilitate the electron transfer and, according to Ferreira *et al.* (2017), stabilize NAD⁺ more effectively after hydride loss than a threonine residue. This shows that the balance between catalytic rate and substrate affinity is delicate and required to achieve optimal enzyme function.

In contrast to the aforementioned variants, DszD_T62R and DszD_T62Y displayed the highest K_m (0.81 mM and 2.07 mM, respectively) values, which is in line with both the volume and positive charge of arginine, as well as the bulkiness of the aromatic tyrosine side chain, that may negatively impact substrate binding (**Figure 12** and **Table 6**). Hence, the size and charge of the residue at this position seems to be critical for enzyme function.

Interestingly, despite having a high k_{cat} (437.6 s⁻¹), DszD_T62Y displayed the highest K_m of all variants, including DszD *wt*. This high k_{cat} may be explained by the results of the *in silico* study performed by Ferreira *et al.* (2017). Their findings may suggest that Tyr62

forms a very weak bond with FMN, the weakest among all variants. This weak bond seems to facilitate the reaction by decreasing the Gibbs energy, even though the bulkiness of the tyrosine side chain doesn't provide an optimal environment for substrate binding. Consequently, when considering both kinetic parameters, the catalytic efficiency of this variant, defined by k_{cat}/K_m , ($212.59 \text{ mM}^{-1} \cdot \text{s}^{-1}$) is lower than that of the *wt* protein ($4933.59 \text{ mM}^{-1} \cdot \text{s}^{-1}$). As a result, the Tyr62 substitution proves to be ineffective in improving the catalytic power of DszD.

When considering the overall catalytic efficiency, the *wt* protein remains the most catalytic efficient version of DszD, which suggests that the native Thr62 provides an optimal balance between substrate affinity and catalytic rate. Nevertheless, substitution by an aspartate residue provides the second highest catalytic efficiency, having the highest k_{cat} . Taken together, our kinetic data highlights the importance of considering both parameters when evaluating enzyme performance. Although the T62D mutation didn't increase the catalytic efficiency of DszD compared to *wt*, it shows promising potential for the optimization of DszD activity. This substitution coupled with others could compensate for the reduced substrate affinity, while keeping a higher k_{cat} , relative to the *wt* protein.

FMN plays a crucial role in regulating the oligomeric states of flavin reductases, inducing structural changes affecting the oligomerization of these enzymes. These changes can impact the rate of the enzymatic activity and the transfer of FMNH_2 to each monooxygenase during the desulfurization process (Aloh et al., 2023). An equilibrium between high oligomeric forms and smaller oligomeric forms seems to be observed, as evidenced by the appearance of a shoulder in the left region of the SEC elution profile (higher molecular weight) and a shift of the main peak toward higher elution volumes (**Figure 13**). However, part of this analysis is biased by the highly flexible N-terminal of DszD predicted by AlphaFold as it may affect the apparent Stokes radius and skew the apparent molecular weight of DszD. For this reason, we decided to delete the first 27 residues of the N-terminal and analyze the truncated version of DszD ($\Delta\text{N-DszD}$) by SEC. N-terminal deletion led to a significant shift in the elution volume and $\Delta\text{N-DszD}$ eluted with an apparent molecular weight of approximately 20kDa consistent with a monomer (**Figure 15** and **Table 9**) that contradicts the dimer prediction in the structural model. A plausible explanation for this conflict may be that the sequence identity between DszD and the proteins used as structural templates is low (~20 %). The amino acid sequences diverge in the dimerization interface with the substitution of several residues, especially those with hydrophobic side chains, which may preclude dimer formation. It is

important to note that the requirement for dimerization was not investigated yet and, therefore, DszD may be functional as a monomer. To understand the oligomeric state and the overall folding of DszD more thoroughly, further studies using dynamic light scattering (DLS) and circular dichroism (CD) spectroscopy should be performed. DLS would provide valuable information about the size and hydrodynamic size and CD would reveal potential changes in the secondary structure in response to either increasing concentrations of FMN or the deletion of the N-terminal.

During our attempts to crystalize DszD *wt*, various conditions yielded crystals (**Figures 16 and 18**) but the X-ray diffraction analysis revealed that they were salt crystals rather than protein crystals. Protein crystallization is a process dependent on multiple factors and is facilitated by protein conformations that are well defined and rigid. Hence, protein conformational flexibility is undesirable as it can interfere with crystal packing and hinder the formation of high-quality diffraction crystals (Deller et al., 2016). To crystallize DszD, we will use the truncated version of DszD (Δ N-DszD) to screen for crystallization conditions (Gräslund et al., 2008).

Based on our results, we conclude that none of the substitutions performed on the residue 62 enhanced the catalytic power of DszD, including DszD_T62D, which was predicted to increase the catalytic activity the most. Further investigation into the structure of DszD, especially on the active site, is necessary to identify other potential residues that could be replaced to effectively improve its catalytic power. Obtaining high-quality diffraction crystals of DszD would provide the detailed structural information needed for such analysis and afford valuable insights on the catalytic mechanism in the active site of both the *wt* protein and variants.

Conclusions and Future Perspectives

Currently, the main employed strategy to remove sulfur from fossil fuels in refineries is HDS. However, this technology is energetically expensive and uses harsh conditions, such as high temperatures and pressures, devaluing the fuel. One promising alternative to HDS that is gaining attention is BDS. This approach is more environmentally friendly and does not require harsh conditions, like HDS. However, the biggest limitation of this technology is the rate at which the microorganisms remove sulfur from the oil. One way to overcome this limitation is to increase the catalytic power of DszD, an enzyme responsible for the rate-limiting step of the 4S pathway, through rational mutagenesis of residues predicted to be relevant for its activity.

In this work, we were able to successfully express and purify both DszD *wt* and variants (T62D, T62L, T62Y and T62R). We performed enzymatic assays to assess whether the variants led to an increase in the enzymatic activity of DszD. Based on our results, we can conclude that none of the variants studied are more catalytically efficient than the *wt*, including DszD_T62D, which was predicted to increase the catalytic power of DszD the most. We also observed an effect of FMN on the oligomerization state and/or in the overall folding of DszD as its concentration increases. However, due to the presence of a highly flexible N-terminal, as predicted by AlphaFold, we couldn't determine the apparent molecular weight of each oligomeric state of DszD. This flexibility may increase the apparent molecular weight of DszD leading to an overestimation of its molecular weight. Further studies using DLS and CD would provide valuable information about the hydrodynamic radius of DszD and possible changes in its secondary structure, respectively. After deleting the first 27 amino-acids of the N-terminal and analyzing this new variant by SEC, we observed that in the absence of the N-terminal DszD is a monomer in solution, which is in conflict with the dimer prediction in the structural model.

Based on the analysis performed by AlphaFold on the full-length 3D structure of DszD, the highly flexible and disordered N-terminal could be inhibiting the crystallization of DszD, as it might hinder crystal packing. To crystallize DszD, we generated a truncated version of the full-length DszD that doesn't feature the disordered N-terminal (the first 27 amino acids) and thus is likely more prone to crystallize. The successful crystallization of DszD could provide great details on the structure of DszD and valuable insights on the catalytic mechanism in the active site of both the *wt* protein and variants.

For future work, further investigation into the structure of DszD is still required, particularly on the active site, to identify potential residues that could be mutated to effectively enhance its catalytic efficiency. Our aim is to find new crystallization conditions that will successfully crystallize the truncated version of DszD to produce high-quality diffracting crystals and determine its 3D structure. Further, we aim to better understand the effect of FMN on the oligomerization state of DszD and its folding using DLS and CD, respectively. Finally, we will investigate whether deletion of the highly flexible N-terminal affects the catalytic power of DszD. These data are crucial to understand how DszD activity can be manipulated to make BSD suitable for industrial use. This could address a major challenge in the petroleum industry and contribute to cleaner fuel production by providing more efficient and cost-effective methods for removing sulfur from crude oil.

Chapter 2: Application of time-resolved serial crystallography on a 4th generation synchrotron light source to dissect the mechanisms of enzyme inhibition and activity at room temperature

INTRODUCTION

Application of time-resolved serial crystallography on a 4th generation synchrotron light source to dissect the mechanisms of enzyme inhibition and activity at room temperature

Proteins are one of the most important classes of molecules in living organisms, being responsible for a wide range of biological functions, including signal transduction, gene transcription, cell death, immune function, structural support, and catalysis of chemical reactions in the cell. It is therefore essential to understand both their structure and function, to understand their role within the cell (Morris et al., 2022).

1.1. Protein Structure Determination

Currently, there is an array of tools and techniques enabling the determination of the structure of proteins. Basic spectroscopic methods, such as circular dichroism or fluorescence spectroscopy can give broad structural information, whereas advanced methods such as X-ray crystallography, cryogenic electron microscopy, and nuclear magnetic resonance provide detailed information at an atomic level of the structure and the dynamics of proteins (Stollar & Smith, 2020).

1.1.1. Protein crystallization

The function of proteins is significantly influenced by their three-dimensional structure and X-ray crystallography has become a core technology to determine the 3D structure of proteins.

Protein crystals are highly ordered arrangements of individual protein molecules. These crystals are composed of repeating unit cells that, when translated into three dimensions, make up the entire crystal lattice. Each unit cell contains an asymmetric unit. The asymmetric unit is the smallest part of the crystal structure to which the symmetry operations can be applied to generate the complete unit cell. The asymmetric unit can contain one or more protein molecules (**Figure 19**) (Krissinel & Henrick, 2007). Upon exposure to an X-ray beam, the protein crystals diffract the X-rays creating a diffraction pattern through constructive and destructive interference (Stollar & Smith, 2020).

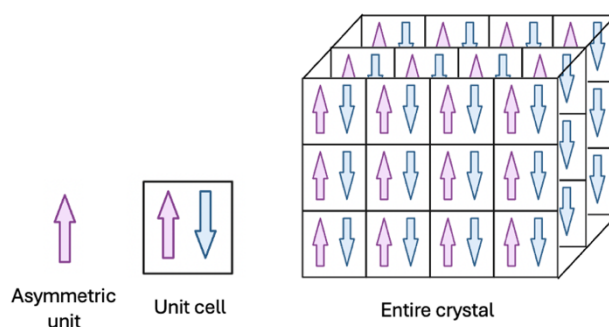


Figure 19. Asymmetric unit, unit cell and entire crystal. The asymmetric unit of a space group consists in the crystallographic unit cell to which symmetry operations can be applied to generate the complete unit cell. The asymmetric unit (purple upward arrow) is rotated 180 degrees to produce a second copy (blue upward arrow). Together the two arrows produce a unit cell. The unit cell is then translationally repeated in three directions to make a crystal. (Image adapted from <https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/biological-assemblies#Anchor-Asym>).

The process of crystallization of a protein is a transition phenomenon that can be effectively captured in a phase diagram that illustrates the various conditions influencing protein solubility and precipitation, as well as the nucleation and the subsequent growth (metastable phase) of the crystals (**Figure 20**) (Shoeman et al., 2023). The nucleation and growth of the crystals can be affected by many factors, including the protein concentration, precipitants, temperature, pH, and nature of the buffer solution. In particular, the type of precipitants (inorganic salts or organic polymers) and their concentration can have a great impact in the protein crystallization process (Liu et al., 2023). Different precipitants have different mechanisms to promote protein crystallization. For example, inorganic salts precipitants (e.g., ammonium sulfate) act by altering the activity coefficient of the water, organic precipitants by altering the dielectric constant of the solvating medium, and in case of high molecular weight polymers like polyethylene glycol (PEG) by increasing molecular crowding (Majeed et al., 2003).

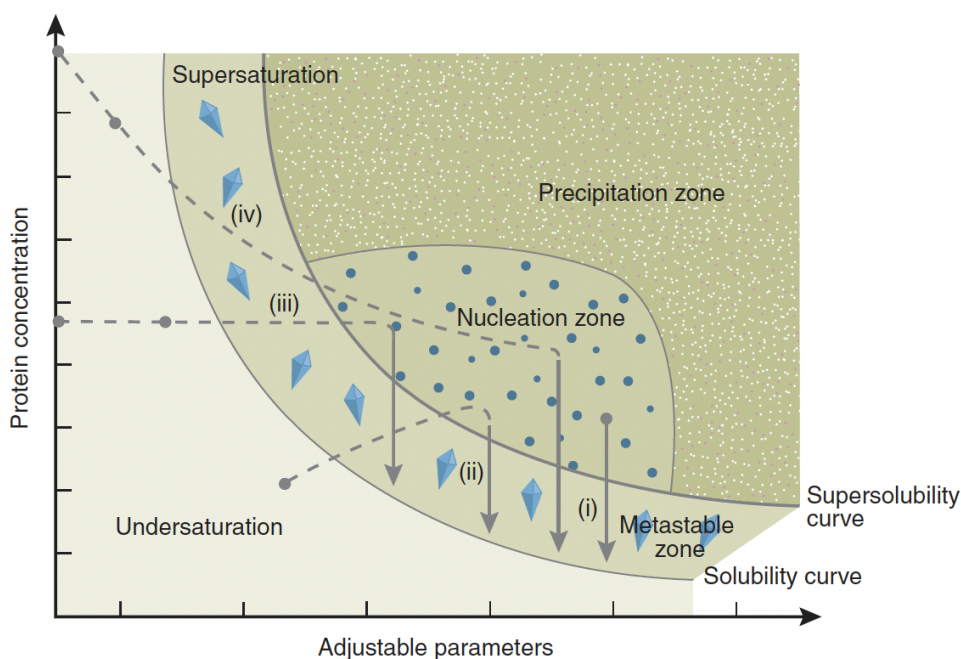


Figure 20. Phase diagram illustrating the different routes to reach supersaturation. The phase diagram is composed of four regions: a region of very high supersaturation, where the protein will precipitate (precipitation zone); a nucleation region, where the supersaturation is moderate; low supersaturation region (below the nucleation region), where the

crystals are stable and might grow (metastable zone); and a undersaturation region, where the protein is soluble and will never crystallize. The four major crystallization methods are represented: (i) microbatch; (ii) vapour-diffusion; (iii) dialysis; and (iv) Free interface diffusion (FID). Figure taken from (Chayen & Saridakis, 2008).

The phase diagram is obtained by altering two parameters simultaneously, therefore representing a two-dimensional “slice” space of the parameters relevant to crystallization. The crystallization process is divided into two stages: the nucleation stage and the growth stage. The nucleation arises from fluctuations in the solution concentration, resulting in reversible and droplet-like molecular clusters that develop into thermodynamically stable crystal nuclei. Subsequently, these nuclei can grow into crystals (Liu et al., 2023). Both nucleation and growth of the crystals depend on the supersaturation of the mother liquor that enables crystallization. Supersaturation acts as a driving force for both processes and determines their occurrence and is illustrated in the phase diagram presented in **Figure 20**. Supersaturation, the initial step in the crystallization process, is a non-equilibrium state in which the quantity of the protein exceeds the solubility limit under certain chemical and physical conditions and the equilibrium is reached by the formation of a solid state (Chayen & Saridakis, 2008). Depending on the level of supersaturation, the phase diagram can be divided into three regions: very high supersaturation, where the protein concentration is well above saturation and an amorphous precipitate form; moderate supersaturation, where both growth and nucleation occur; and lower supersaturation, where only growth occurs (**Figure 20**). The boundaries between these regions are not well defined, as they depend on kinetic factors that can vary from specific experimental conditions (Russo Krauss et al., 2013). The supersaturation state not only has a major importance in the nucleation and growth of the crystals, but also greatly affects the morphology, quality, and size of the crystals (Liu et al., 2023).

1.1.2. Protein crystallization methods

To achieve the supersaturation state, it is important to have a solution in which the concentration of the protein is below its solubility limit. This can be achieved in several ways: (i) changing the protein itself by altering, for example, the ionization state through changes in the pH; (ii) altering the chemical activity of the water, by adding salt; (iii) changing the degree of attraction of one protein molecule to another, by changing the pH; or (iv) altering the interactions of protein molecules with the solvent with the addition of polymers or ions (Mcpherson, 2004). The different crystallization methods (e.g., vapor diffusion crystallization; batch crystallization, seeding, etc.) take different paths through the phase diagram to reach this supersaturation state.

Vapor diffusion is one of the most widely used methods for the screening of optimal protein crystallization conditions. In this method, the protein and the crystallizing agents are combined usually in a 1:1 ratio, within a drop that can be hanging on top of (hanging drop) or sitting above (sitting drop) the reservoir, known as mother liquor (**Figure 21**), against which the protein solution is equilibrated. In the vapor diffusion method, water moves from the drop containing the protein solution and the crystallizing agents to the reservoir solution, concentrating the protein and the precipitant (Pusey & Ayygün, 2017).

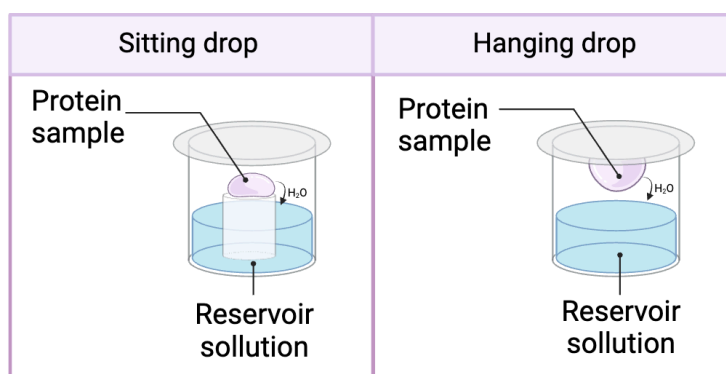


Figure 21. Vapor diffusion method. Experimental setup for the hanging drop and sitting drop vapor diffusion methods. On the right-hand side is the setup of a hanging drop crystallization plate, where the protein drop is suspended above the reservoir solution from a seal and does not contact the crystallization plate. In the sitting drop setup, displayed on the left-hand side, the protein drop sits directly on a surface of the crystallization plate above the well containing the reservoir solution. (Pusey & Ayygün, 2017).

In the batch method, the protein and the reservoir solution are simply mixed together in a vial (Pusey & Ayygün, 2017). In this method, the nucleation zone is immediately reached shortly after mixing protein and reservoir solutions, unlike with a vapor diffusion setup. The combination should result in a supersaturated solution, causing the protein to nucleate more rapidly. The batch method allows the nucleation zone to be more exploited, resulting in more nucleation (Beale et al., 2019).

Seeding is a technique that promotes the nucleation and growth of the crystals, where small pieces of crystals (macroseeding) or submicroscopic crystals (microseeding) are introduced into a solution containing the protein of interest. This can be done in both vapor diffusing and batch methods. The seed suspensions can be prepared by harvesting crystals and finely crushing them with a needle or a seed bead before diluting them into the mother liquor, creating a seed stock. This results in a non-homogeneous seed solution with different seeds sizes, which can then carry over the growth of the crystals (Shoeman et al., 2023).

Given their biological relevance, another important focus of protein crystallography is the determination of protein-ligand complexes. There are several methods to obtain crystals of protein-ligand complexes, being the most used soaking and co-crystallization (Rondeau & Schreuder, 2015). In co-crystallization experiments, the ligand and the protein are mixed together in a solution before the formation of the crystals and the crystals grow with the ligand bound to the protein of interest. However, because crystal growth can take several days, the success of co-crystallization can be difficult to estimate (Hassell et al., 2007). In the soaking method, a solution containing the ligand is added to a drop already containing protein crystal(s) (**Figure 22**). During the protein crystallization process, water and other solvent molecules are frequently trapped inside of the protein lattice forming solvent channels, that can account for 25% to 75% of the volume of the protein crystal, depending on the crystal packing (Sprenger et al., 2021). These channels allow the diffusion of low molecular weight ligands, co-factors, and substrates to the protein binding sites. The binding of ligands can cause small conformational changes in the crystal lattice, without damaging the structure of the crystal (Rondeau & Schreuder, 2015). The diffusion of the ligands into the binding site in the protein can be fast, taking only seconds to minutes but it can also take hours to populate the binding pocket of the protein. For successful soaking of the crystal, not only the affinity of the ligand is important, but also its diffusion kinetics through the crystal channels (Wienen-Schmidt et al., 2021).

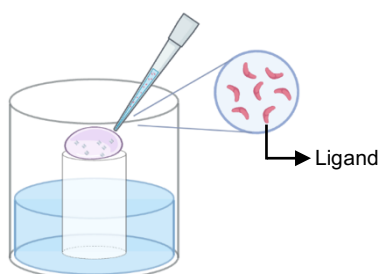


Figure 22. Experimental setup for soaking protein crystals in a sitting drop. After crystal growth, the ligand solution is added to the drop. Different conditions should be tested to guarantee that the protein crystals remain intact and ligand diffusion towards the binding site.

1.2. X-ray Crystallography

X-ray crystallography is a widely used technique for determining the 3D structure of biomolecules, such as proteins. To determine the structure of a protein using this technique, a protein crystal is exposed to X-rays with a wavelength of the order of the atomic spacing or smaller. After interacting with the electron clouds surrounding the atoms in the crystal, the X-rays are scattered in different directions and intensities,

depending on the electron concentration and whether the scattered X-rays interfere constructively or destructively (Cowtan, 2003). Because of the regular and repeating arrangement of atoms in the crystal, the scattered X-rays create a diffraction pattern that contains all the information about the positions of the atoms in the protein, and this pattern can be used to deduce the original structure of the protein using the Bragg's Law (H. Wang & Wang, 2017) (Brito & Archer, 2020). However, in addition to knowing the intensities of the diffraction spots in the pattern, it is necessary to know the phases of the spots. To solve this problem, the phase information can be obtained by computational means, such as molecular replacement (MR), single/multi-wavelength anomalous diffraction (SAD/MAD), etc. Once the intensities and phases of the spots are known, the protein structure can be determined (Brito & Archer, 2020).

During the data collection process, the crystals are exposed to X-ray radiation, and the nature of this radiation is damaging to the crystals, which is a limiting factor in this technique. To mitigate this problem, the crystals are kept at cryogenic temperatures, as radiation damage accumulates more slowly at low temperatures (Botha et al., 2015). In cryocrystallography, protein crystals are soaked in a cryoprotectant solution to prevent the water in solvent channels from forming crystalline ice. These cryogenic temperatures allow the collection of a complete data set from a single crystal (Ilari & Savino, 2008). In this technique, a single crystal is cooled to cryogenic temperatures and rotated in multiple orientations while being exposed to an X-ray beam to collect a complete data set.

1.3. Time-resolved serial crystallography

Proteins are dynamic systems, and their structures fluctuate between numerous conformational sub-states. To study their function, it is important to understand the interplay between their structure and dynamics (Šrajer & Schmidt, 2017). Cryo-crystallography has been successful in obtaining high-resolution structures at temperatures of 100 K but it fails to capture the structural dynamic changes of proteins that occur at physiological temperatures (Botha et al., 2015). Moreover, collecting crystallographic data at room temperature enables time-resolved experiments and bypasses the requirement for cryoprotectants. For these reasons, room-temperature data collection is being more heavily used to gather relevant structural information that could not be otherwise obtained in cryogenic conditions.

Time-resolved serial crystallography overcomes some of the issues inherent to cryogenic crystallography and provides unique insights into protein dynamics at room temperature, such as information about protein catalysis, allostery, oxidation states, side chain motions, isomerization reactions, and molecular binding events (Schulz et al.,

2022). The changes in atomic coordinates within the protein as a reaction proceeds, can be studied in detail using time-resolved crystallography. This technique preserves the natural state, dynamic behavior, and structural flexibility of the protein of interest (Meszaros & Westenhoff, 2024) and provides atomic-level information of the interaction between proteins and their ligands (Sharma et al., 2023). Thus, the goal of time-resolved crystallography is to determine both the chemical and kinetic mechanisms with detailed characterization of intermediate states and the reaction rate constants (Šrajer & Schmidt, 2017). In serial crystallography, the diffraction data is collected in a serial fashion and the data from many crystals are then merged together to obtain a complete data set (**Figure 23**) (Botha et al., 2015). This technique relies heavily on the use of advanced light sources, such as a combination of X-ray free-electron laser (XFEL) sources or a third or fourth generation synchrotron (SSX). These light sources can produce extremely bright, focused, and ultrafast X-ray pulses, allowing rapid collection of data from very small crystals (microcrystals or nanocrystals) before the radiation has time to destroy the sample, known as the "diffraction before destruction" principle. The emergence of these extreme bright X-ray sources allowed serial crystallography experiments in a wide range of time resolutions, ranging from femtoseconds for XFEL to milliseconds for monochromatic synchrotron sources (Pearson & Mehrabi, 2020).

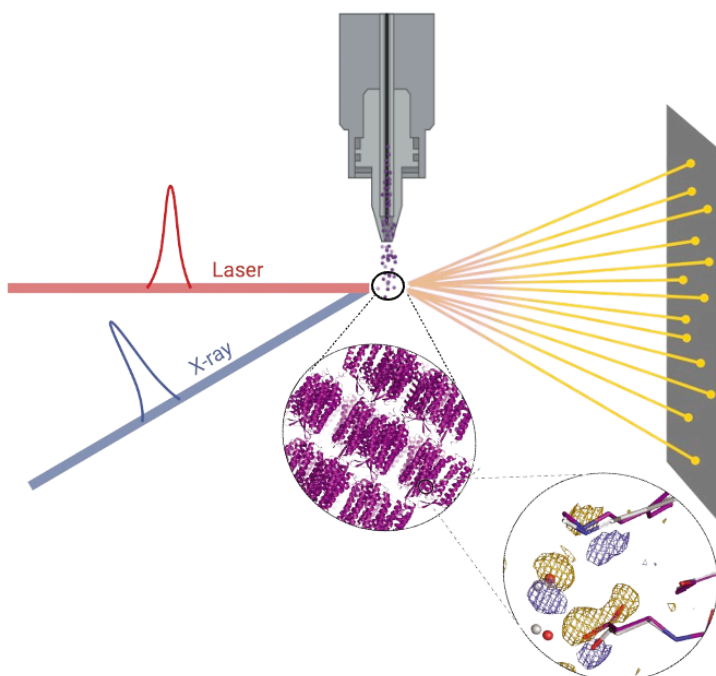


Figure 23. Schematic illustration of a TR-SSX. In this illustration a stream of microcrystals (purple) is injected across the X-ray beam (blue) at a synchrotron or XFEL source, immediately after being activated by a laser pulse (red). Figure adapted from (Brändén & Neutze, 2021).

1.3.1. Sample requirements

In time-resolved serial crystallography, crystal size and homogeneity are crucial factors that can significantly affect the quality and success of data collection. In addition, the size of the crystals depends also on the requirements of different sample delivery and reaction initiation methods, which must be tailored for the planned experiment. For example, while fixed target and tape-drive delivery methods are relatively flexible in size requirements, liquid jet methods require smaller crystals, typically smaller than 10 μm . The penetration depth of the optical triggers and substrate diffusion time are also affected by the crystal size. Furthermore, homogeneity of the crystals is critical to ensure that the reaction initiation is similar for all crystals (Schulz et al., 2022).

1.3.2. Obtaining microcrystals

The crystallization conditions used to obtain single crystals are a good starting point for producing smaller crystals that can be used in serial-crystallography experiments. Typically, these conditions can be adjusted to enhance nucleation, resulting in a drop with higher density of smaller crystals (Schulz et al., 2022). Vapor diffusion methods are often used to grow microcrystals, mostly by increasing the concentrations of protein and precipitant. However, time-resolved serial crystallography requires a large amount of microcrystals (in the order of thousands), so vapor diffusion might not be the ideal method since it is time consuming and pooling together the drops may be detrimental to the crystals. Consequently, the batch method is normally preferred to obtain a larger volume of microcrystals (more than 500 μL) because it enables fast nucleation and therefore the resulting crystals usually have a homogeneous size distribution (Schulz et al., 2022). In addition, by combining batch with micro-seeding, it is often possible to obtain high yields of microcrystals of the desired size in a highly homogeneous solution (Shoeman et al., 2023).

1.3.3. Sample delivery

The fast development of sample delivery methods has paralleled the advancements in extreme light sources for serial crystallography, enabling the fast sample replacement that is necessary for serial crystallography experiments (Monteiro et al., 2021). Both time-resolved serial synchrotron crystallography (TR-SSX) and time resolved femtosecond crystallography require sample delivery methods that can deliver thousands of crystals into the X-ray beam (SFX), while allowing reaction initiation. These sample delivery platforms include fixed targets, enclosed microfluidics devices, tape drives and liquid and viscous jets (**Figure 24**) (Pearson & Mehrabi, 2020)

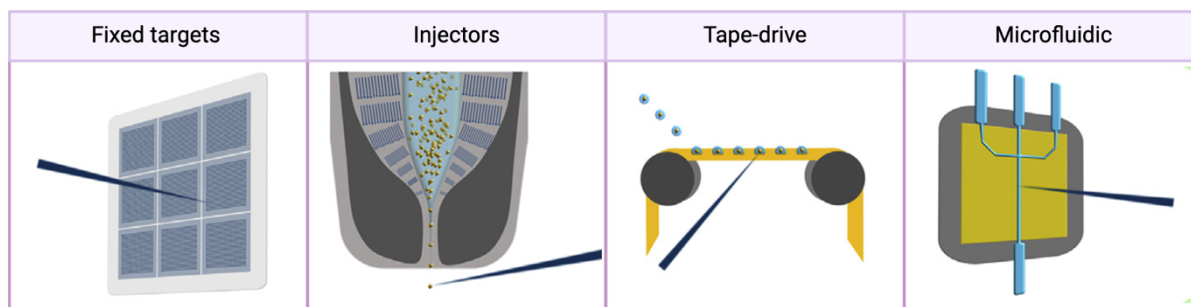


Figure 24. Schematic illustration of sample delivery methods for TR-SSX. In this illustration starting from the left: fixed target, injectors, tape-drive and microfluidic devices. Figure adapted from (Pearson & Mehrabi, 2020).

Fixed targets

In fixed target experiments, the sample is immobilized on a solid support, which is then scanned by the X-ray beam. An advantage of this sample delivery method is that the geometry and crystal distribution can be arranged so that every crystal is probed (Grünbein & Nass Kovacs, 2019). In addition, the crystals can be grown directly on the solid support, reducing the damage to the crystals during harvesting (Pearson & Mehrabi, 2020).

Liquid and viscous injectors

In this sample delivery method, microcrystals are suspended in their mother liquor or in a suitable carrier medium and delivered into the X-ray interaction region of a XFEL or synchrotron in a continuous jet or in droplets. The jet of microcrystals is then synchronized with an X-ray pulse. Over the years, different injectors to deliver samples with different viscosities have been developed, namely liquid and viscous injectors (Grünbein & Nass Kovacs, 2019).

The advantage of using injectors in time-resolved serial crystallography experiments is that it is far easier to keep the crystals stable and hydrated while embedded in a suitable matrix. In addition, these injectors can be used for “mix-and-inject” experiments, where the microcrystals slurry can be mixed with the ligand immediately before being delivered into the X-ray beam region (Hadian-Jazi et al., 2021).

Tape-drive

In the tape-drive method, microcrystals are injected from an injector and placed on a polyimide tape, which transports the crystals to the X-ray interaction region and also enables ligand mixing experiments (Nam, 2024).

Microfluidics devices

Microfluidic devices use multiple channels controlled by syringe pumps where the microcrystals and the ligands can be mixed before being delivered to the X-ray beam (Schulz et al., 2022).

1.3.4. Reaction initiation

Despite significant advances in X-ray sources and sample delivery methods, one of the major challenges in time-resolved structural biology is how reactions are triggered. To obtain clear visualization of structural changes, the macromolecules in the crystal or solution samples must be synchronized, which requires the reaction to be initiated uniformly throughout the sample on a time scale that matches the relevant reaction steps, ensuring the interpretation of the time-dependent electron-density maps (Šrajer & Schmidt, 2017).

All time-resolved serial crystallography experiments use the pump-probe approach to initiate a reaction. In this approach, the reaction is initiated in a crystal by the “pump”, which is then probed by X-rays after a defined time delay (**Figure 23**). Several methods have been employed to trigger reactions, including excitation by intense laser pulses, diffuse mixing, light activation of caged substrates, and temperature jumps (Pandey et al., 2020).

Most of these experiments have focused on the study of naturally photoactivated molecules, such as chromophores that can undergo *cis-trans* isomerization (e.g., rhodopsins, photoactivatable yellow proteins, and green fluorescent protein) or photon-induced conformational changes (e.g., photosynthetic reaction center and photosystem II). Experiments with these proteins have the advantage that their activation is very fast (femtosecond to nanosecond range) and efficient, and the speed of activation allows to achieve high time resolution (Poddar et al., 2022). However, only a small fraction of proteins (less than 0.5% of proteins) is naturally photoactivable and alternative approaches have been developed. One of these approaches involves using infrared pulses to generate temperature jumps, which help to overcome thermal activation barriers within nanoseconds, since the energy is rapidly transferred from water to proteins (Schulz et al., 2022). For slow reactions, where the desired time resolution is in the range of millisecond to second, diffuse mixing is a good approach. In the diffuse mixing setup, the reaction starts by mixing the crystals with the ligand and, after a specific time delay, the mixture is exposed to X-rays (Pandey et al., 2020). The reaction can be followed on different time scales by varying the time delay. Nonetheless, when the reaction time scales are faster than the diffusion of the substrate, light activation remains

the viable option. The use of photocages is a promising approach for implementing light activation, with time scales longer than 100 ms (Monteiro et al., 2021).

Photocages are molecules, such as proteins or their ligands, to which a photoactivatable protective group is covalently attached, thereby inactivating the system. The system is activated by a pulse of light that cleaves the photocage group and releases the active molecule (Figure 25). One of the factors to consider when using photocages, is that the cleavage rate of a photocage group cannot be slower than the reaction rate of the experiment, since the slowest event will limit the time resolution of the experiment. The electron density maps are then calculated at different time delays revealing the structural changes that occur during the light-induced reaction (Monteiro et al., 2021).

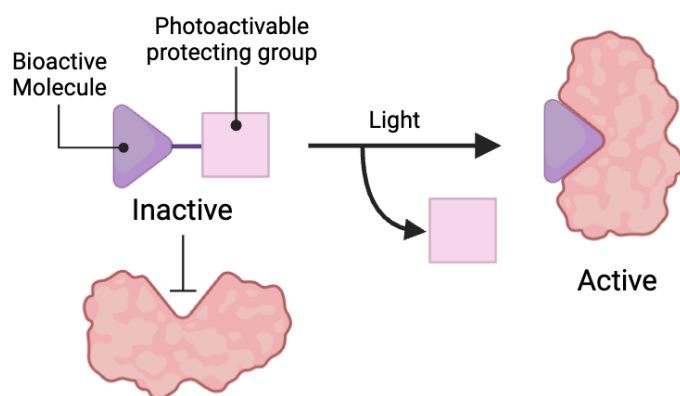


Figure 25. Schematic diagram of a photocaging reaction. The system is activated by a pulse of light that cleaves the photocage group and releases the active molecule, enabling the reaction.

1.3.5. Importance of optimizing protocols for time-resolved serial crystallography

Time-resolved serial crystallography represents a significant advancement in structural biology, enabling the capture of molecular processes of proteins in real time with atomic resolution. However, this technique depends heavily on the optimization of protocols to ensure data quality, experimental efficiency, and the ability to capture dynamic processes. In addition, establishment of standardized protocols is crucial for ensuring reproducibility in experiments. Reproducibility and standardization are important features of this technique, as slight variations can lead to significant differences in the structural data obtained (Schulz et al., 2022).

Protocols provide guidelines for optimizing various aspects of the experimental set-up, such as sample preparation, sample delivery methods and data acquisition strategies. For instance, optimizing protocols for microcrystal production, the choice of sample

delivery methods, and the method of initiating reaction can greatly influence the quality of the data collected (Schulz et al., 2022).

The aim of this project is to optimize protocols for time-resolved serial crystallography in the new ID29 beamline of the recently upgraded European Synchrotron Radiation Facility - Extremely Brilliant Source (ESRF-EBS) and to apply them to the study of the proteolytic mechanism of human thrombin, as well as to the mode of inhibition of this protease by specific natural anticoagulants from hematophagous animals. However, commercially purified thrombin is very expensive, so these protocols are to be initially optimized using a cysteine desulfurase, IscS, that was previously purified, crystallized, and structurally characterized in the lab. Moreover, a photocaged version of IscS's substrate is commercially available, enabling time-resolved experiments.

1.3.5.1 Cysteine desulfurase, IscS

IscS is a cysteine desulfurase enzyme that belongs to an iron-sulfur cluster biosynthetic pathway found in many bacteria and in the mitochondria of eukaryotes. In this pathway, IscS converts the substrate L-cysteine (L-cys) into L-alanine, delivering the atom of sulfur to IscU, a scaffolding protein for iron-sulfur cluster assembly (**Figure 26**). IscS is a homodimer with each subunit containing a 5'-pyridoxal phosphate (PLP) cofactor and a mobile S-transfer loop cysteine residue (Lin et al., 2020).

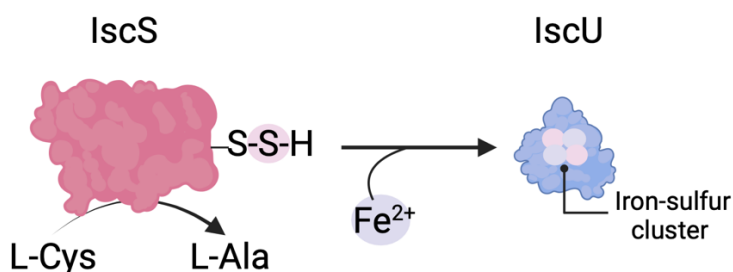


Figure 26. Simplified illustration of the enzymatic mechanism of IscS. IscS converts L-Cys into L-Ala and provides the sulfur to IscU for iron-sulfur cluster assembly.

For time-resolved experiments with IscS, we will use a synthetic commercially available substrate for this enzyme, S-(2-nitrophenyl)-L-cysteine (**Figure 27**). This substrate incorporates a photocaged moiety (o-nitrobenzyl) covalently linked to the cysteine residue, inhibiting its reactivity with IscS upon mixing. Subsequently, a pulse of light at a specific wavelength will be applied to the sample, triggering the photolytic cleavage of the photocaged group. This activation will enable the substrate to be processed by IscS, allowing the subsequent reaction dynamics to be followed with time-resolved serial

crystallography after a time-delay. This approach will facilitate the detailed observation of transient states during the enzymatic reaction, providing valuable insights into the mechanistic pathways of IscS.

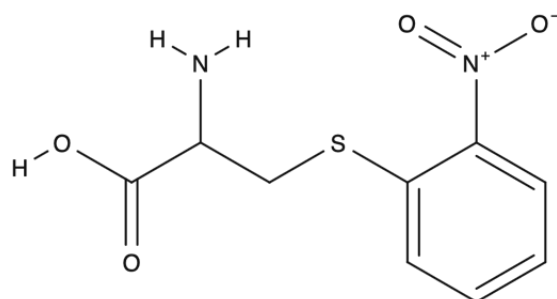


Figure 27. S-(2-nitrophenyl)-L-Cysteine structure.

Motivation and Aim

Protein-protein interactions are at the core of many biological processes. These interactions play a key role in the regulation of biochemical pathways, signaling cascades, and play a fundamental role in substrate recognition. Moreover, proteins are dynamic systems, with their structures fluctuating between numerous conformational substates. To fully grasp their functionality, it is important to understand the interplay between their structure and dynamics.

Time-resolved serial crystallography is an advanced experimental technique that allows the determination of the structure of biological molecules at an atomic level and investigate the dynamics of protein-protein interactions, substrate recognition and enzymatic processing. However, the effectiveness of time-resolved crystallography relies on the optimization of protocols that ensure data quality and improve experimental efficiency to capture highly dynamic processes. It is crucial to ensure reproducibility and standardization, as even minor variations can impact the structural data obtained.

With this project, we aim to optimize protocols for time-resolved serial crystallography at the new ID29 beamline of the European Synchrotron Radiation Facility - Extremely Brilliant Source (ESRF-EBS). As a proof of concept, we will develop these protocols to investigate the proteolytic mechanism of human thrombin and the mode of inhibition of this protease by specific natural anticoagulants from hematophagous animals. This decision was based in the strong expertise that the research group has developed through years of working with thrombin and the detailed information already available about its mechanism. However, due to the high cost of commercially purified thrombin, we will first optimize these protocols using IscS that can be easily produced in our lab and its 3D structure has been already determined. Furthermore, we have previously established crystallization conditions for IscS and a photocaged substrate is commercially available.

With the present work we aimed to optimize protocols for the preparation of microcrystals of IscS that can be mixed with photoactivatable substrates in pump-probe experiments to reveal the molecular details of substrate recognition and the hydrolysis cycle. For this, we will determine the corresponding crystallographic structures at room temperature in the ID29 beamline of ESRF (Grenoble, France), which is dedicated to serial and time-resolved crystallography experiments using a 4th generation synchrotron light source.

MATERIALS AND METHODS

2.1. Materials

2.1.1. Bacterial Strains

Table 10. Bacterial strains used in the following protocols. Bacterial strains, their genotype and application in the following protocols.

Bacterial strains	Genotype	Application	Methods section
<i>E. coli</i> DH5 α	<i>F</i> - <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96</i> Φ 80 <i>dlacZ</i> Δ M15 <i>deoR</i> Δ (<i>lacZ</i> YA- <i>argF</i>) <i>U169 nupG</i> , <i>hsdR17(rK- mK+)</i> , λ	General cloning and sub-cloning	2.2.1
<i>E. coli</i> BL21 (DE3)	<i>F</i> ⁻ <i>ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 *lacI</i> <i>lacUV5-T7 gene 1 ind1 sam7 nin5)</i>	Protein expression	2.2.2

2.1.2. Cell Culture Media

Table 11. Cell culture media used in the following protocols.

Medium	Composition	Methods Section
Lysogeny Broth medium (LB)	1 % (w/v) Tryptone	2.2.1
	0.5 % (w/v) Yeast Extract	
	1 % (w/v) NaCl	2.2.2
Lysogeny Broth medium plates (LB agar)	1 % (w/v) Tryptone	2.2.1
	0.5 % (w/v) Yeast Extract	
	1 % (w/v) NaCl	2.2.2
	1.5 % (w/v) Agar	

2.1.3 Buffer Solutions

Table 12. Buffers used in the following protocols.

Buffer	Composition	Methods section
High salt buffer	20 mM Tris pH 8.0	2.2.3
	1 M NaCl	
	10 mM Imidazole	
	20 mM β -mercaptoethanol*	
Low salt buffer	20 mM Tris pH 8.0	2.2.3
	150 mM NaCl	
	10 mM Imidazole	
	20 mM β -mercaptoethanol*	
Elution Buffer	20 mM Tris pH 8.0	2.2.3
	150 mM NaCl	
	500 mM Imidazole	
	20 mM β -mercaptoethanol*	
Desalting Buffer	20 mM Tris-HCl pH 6.8	2.2.3
	150 mM NaCl	
	20 mM β -mercaptoethanol*	
Loading Buffer	200 mM Tris-HCl pH 6.8	2.2.4
	20% (v/v) Glycerol	
	10% (v/v) SDS	
Running Buffer	25 mM Tris-HCl pH 6.8	2.2.4
	192 mM Glycine	
	0.1% (v/v) SDS	

*added fresh in the purification day

2.2. Methods

2.2.1. DNA replication

E. coli DH5 α (**Table 10**) cells were transformed by adding 100 ng of plasmid and incubated on ice for 15'. The cells were then heat shocked at 42 °C for 30" and cooled down on ice for 5'. LB medium (600 μ L) (**Table 11**) was added to the cells, which were incubated at 37 °C for 60' with agitation of 180 rpm (IKA[®] KS 4000 ic) and subsequently plated on LB agar plates (**Table 11**) supplemented with 50 μ g/mL kanamycin (KAN). The plates were incubated at 37 °C and 180 rpm (IKA[®] KS 4000 ic) overnight. The next day, one isolated colony was used to inoculate 10 mL of LB medium supplemented with KAN (50 μ g/mL). The pre-culture was incubated overnight at 37 °C and 180 rpm (IKA[®] KS 4000 ic). In the following day, the DNA was extracted and purified using the NZYMiniprep Kit (NZYTech), according to the manufacturer's protocol. The final plasmid concentration and its purity were assessed by measuring the absorbance at 260 nm and the A260nm/A280nm ratio, respectively, using a NanoDrop One (Thermo Fisher Scientific) spectrophotometer. The plasmids were stored at -20 °C.

2.2.2. Protein Expression

E. coli BL21 (DE3) cells (**Table 10**) were transformed following the same protocol used for DNA replication in *E. coli* DH5 α cells. The following day, one isolated colony was used to inoculate 200 mL of LB medium (**Table 11**) supplemented with 50 μ g/mL KAN and incubated at 37 °C and 180 rpm (IKA[®] KS 4000 ic) overnight.

The following day, 20 mL of the pre-culture were added to 500 mL of LB medium (**Table 11**) supplemented with 50 μ g/mL KAN in 2 L flasks. The flasks were incubated at 37 °C at 180 rpm (IKA[®] KS 4000 ic), until the OD at 600 nm reached 0.8. At this point, 1 mM IPTG was added to induce protein expression. The cells were incubated for 4 hours at 37 °C and 180 rpm (IKA[®] KS 4000 ic). Then, the cells were harvested by centrifugation in a high-speed centrifuge at 4000 rpm for 20' at 4 °C (Beckman[®] Avanti J-26 XP, JLA-81000 rotor). The supernatant was discarded, and the pellet was resuspended in 3x the wet weight of the cells in volume of cold lysis buffer (**Table 12**) and stored at -20 °C.

2.2.3. Protein Purification

IMAC was used to purify IscS fused with a C-terminal 6xHis affinity tag and a S-peptide solubility tag (S-tag). The cell suspensions were initially thawed at room temperature and then sonicated in a Sonoplus HD 2200/HD 3200 using a microtip MS73 (BANDELIN) for a total of 5' (30" ON; 30" OFF; 30% power). Afterwards, the lysate was clarified by centrifugation at 17000 rpm at 4 °C for 40' (Beckman[®] Avanti J-26 XP, JA-25.50 rotor). The supernatant was filtered with a low binding 0.45 μ m PVDF Filter (CHROMAFIL[®] Xtra PVDF-20-25) before loading onto a Ni²⁺-charged HisTrap[™] IMAC HP 5 mL column (Cytiva) using a peristaltic pump (Bio Rad[®] Econo Gradient). The column was previously washed with 5 CV of water and equilibrated with 5 CV of low salt buffer (**Table 12**). The clarified lysate was recirculated in the column for 40' at 2.5 mL/min and the FT was then collected. The column was washed with 100 mL of low salt buffer, followed by 100 mL of high salt buffer, 50 mL of low salt buffer and 50 mL of elution buffer (**Table 12**). For each purification step, a sample was taken for SDS-PAGE analysis. After this analysis, IscS_S-His was eluted in the low salt buffer containing 10 mM of imidazole. To this fraction, TEV protease was added to the solution in a 1:10 molar ratio relative to IscS_S-His, and the mix was left overnight rotating at 4 °C.

The following day, the column was equilibrated with 5 CV of low salt buffer using a peristaltic pump (Bio Rad[®] Econo Gradient) and the protein solution was recirculated for 40' at 2.5 mL/min. IscS was collected in the FT. The column was then washed with 100 mL of low-salt buffer and 50 mL of elution buffer (**Table 12**). In each step of

purification, a sample was taken for SDS-PAGE gel analysis. Afterwards, the FT was concentrated using a 30 kDa cut-off ultrafiltration device (Amicon® Ultra-15). The protein concentration was measured on a NanoDrop One spectrophotometer (Thermo Fischer Scientific) using the theoretical molar extinction coefficient of IscS at 280 nm ($\epsilon_{280\text{nm}}$ (1mg/mL) = 0.918). Then, for both sample clean up and quality control, the sample was injected on a size exclusion chromatography (SEC) column (Superdex75 10/300 GL, Cytiva). The SEC column was equilibrated with 20 mM Tris pH 8; 150 mM NaCl; 20 mM β -mercaptoethanol on a AKTAPure™ (Cytiva) before sample injection (**Table 12**). Fractions containing IscS with the highest levels of IscS were pooled together and concentrated using a 10 kDa cut-off ultrafiltration device (Amicon® Ultra-0.5). Protein concentration was measured on a NanoDrop One spectrophotometer (Thermo Fischer Scientific) using the theoretical molar extinction coefficient of IscS at 280 nm ($\epsilon_{280\text{nm}}$ (1mg/mL) = 0.918).

2.2.4. SDS-PAGE analysis

All expression and purification samples were analyzed on 12.5% (w/v) acrylamide resolving gels with 4% (w/v) of acrylamide stacking gels, since the molecular weight of all purified proteins was comprised between 20 and 80 kDa. To directly compare the level of protein expression, the same amount of cells were collected before (T0) and after the 4h of protein induction with IPTG (T4). The samples were centrifuged at 4000 rpm (Hitachi CT15RE centrifuge) for 2' at room temperature. The pellet was resuspended in 100 μ L of Buffer U (50 mM Tris pH 7.5, 100 mM NaCl, 6 M urea) and loading buffer (2% (w/v) SDS, 5% (v/v) glycerol, 50 mM Tris-HCl pH 6.8, 0.01% (w/v) bromophenol blue, 0.01% (w/v) 2-mercaptoethanol) was added to the samples. The gels were assembled into a Mini-PROTEAN® Tetra Cell System (Bio-Rad®) and ran at 200 V in running buffer (25 mM Tris-HCl pH 8.3, 192 mM Glycine, 0.1% (w/v) SDS) (**Table 12**) at room temperature, for approximately 45', stained with ProBlue Safe Stain (GiottoBiotech), and imaged in a Chemidoc Imaging System (Bio-Rad®).

2.2.5. Crystallization

An initial screening was carried out using crystallization solutions from two crystallization kits: MIDAS™ and JCSG+ (Molecular Dimensions). Sitting drop plates for vapor diffusion were set up with a protein/precipitant ratio of 1:1 in drops of 2 μ L. The 24-well crystallization plates (Crychem plates, Hampton Research) were incubated at 20 °C and crystals appeared after 2 days. Two sequential rounds of seeding were performed, in which different seed dilutions were tested to produce homogeneous microcrystals (**Figure 29**). Seeding was performed by first crushing the crystals with a Seed Bead™

(Hampton Research) according to manufacturer's classical protocol to obtain a concentrated seed stock. Serial 1:10 dilutions were made from the concentrated seed stock. The drop contained 3 μL purified IscS (16 mg/mL), 2 μL reservoir solution and one μL seed stock. Crystallization plates were kept at 20 °C. Crystals appeared 24 hours after seeding. The crystallization plates were incubated at 20 °C and were checked after 24 hours, two days, three days, one week, and then weekly for a month.

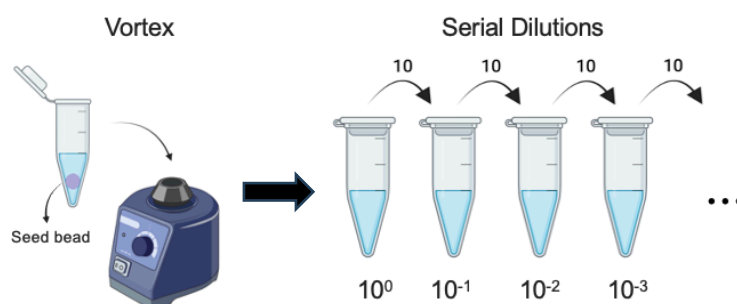


Figure 28. Seed preparation scheme using Seed Bead™ (Hampton Research protocol). The crystals are first crushed with a Seed Bead by vortexing. Different seed stocks are prepared through 1:10 serial dilutions of the initial seed stock preparation.

RESULTS

3.1. Expression and purification of IscS

IscS protein was previously cloned into a pET30 (+) plasmid under the control of the IPTG-inducible T7 promotor and were expressed in *E. coli* BL21 (DE3) cells. The cells were incubated at two different temperatures, 30 °C and 37 °C to test which yielded higher levels of protein expression. Samples were collected at different time points, before IPTG induction (T0) and 4 hours after IPTG induction (T4). Both these samples were analyzed by SDS-PAGE. Expression of IscS-S-His was confirmed by the appearance of a band with the expected molecular weight $\approx$ 47 kDa (S-tag molecular weight is 1.8 kDa and IscS has 45 kDa) after IPTG induction, and a higher protein yield was observed at 37 °C compared to 30 °C (**Figure 29**). Based on this results, subsequent protein expressions were conducted at 37 °C to maximize protein yield, since it's particularly important for crystallization studies, which require high protein concentrations.

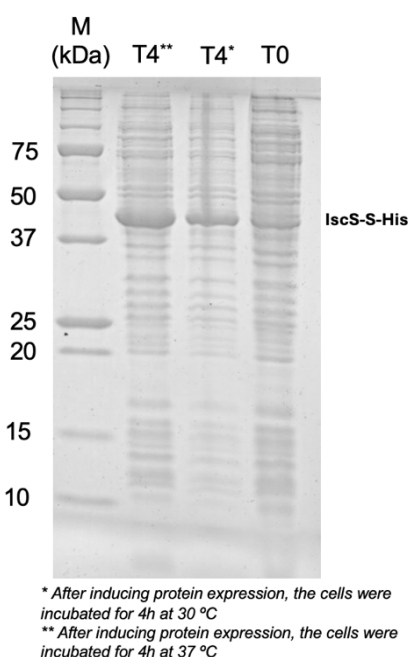


Figure 29. SDS-PAGE gel of IscS protein expression. SDS-PAGE analysis of the fractions collected before (T0) and 4 hours after IPTG induction (T4). The protein expression of IscS was confirmed by the appearance of a band with the expected molecular weight of 47 kDa after IPTG induction at the T4 lanes. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad).

After protein expression, IscS was purified in two sequential IMAC steps separated by an overnight incubation with TEV protease. IscS was fused to a S-peptide solubility tag (S-tag) and a histidine affinity tag (6xHis-tag). The S-tag was included to enhance solubility, while the 6xHis-tag facilitated purification via IMAC using Ni^{2+} ions. A TEV

protease recognition sequence (ENLYFQ▼S) was inserted between the solubility tag and the IscS sequence. This insertion allowed for the removal of both tags through cleavage with TEV protease, resulting in tag-free IscS.

The IMAC purification steps had been previously optimized in the laboratory. The first purification step was designed to isolate IscS_S_His from other soluble cellular components and involved several steps, including washing with low salt and high salt buffers followed by an elution step (**Table 12**). SDS-PAGE analysis was performed on fractions collected during each purification step. IscS-S-His eluted predominantly in the low salt wash with 10 mM imidazole, although a minor amount was also observed in the final elution step with 500 mM imidazole (**Figure 30a**). The protein bands observed on the gel correspond to the predicted molecular weight of IscS-S-His (47 kDa). To cleave the affinity and solubility tag, TEV protease was added directly to the low salt wash fraction and incubated overnight. After the overnight cleavage, a second purification step was performed to separate IscS from the affinity and solubility tag, as well from the TEV protease and the uncut protein (IscS-S-His) that bound to the IMAC column and were eluted with 500 mM imidazole. IscS, lacking the His-S-tag, was collected in the flow-through fraction. SDS-PAGE analysis confirmed the presence of IscS at its predicted molecular weight of 45 kDa (**Figure 30b**).

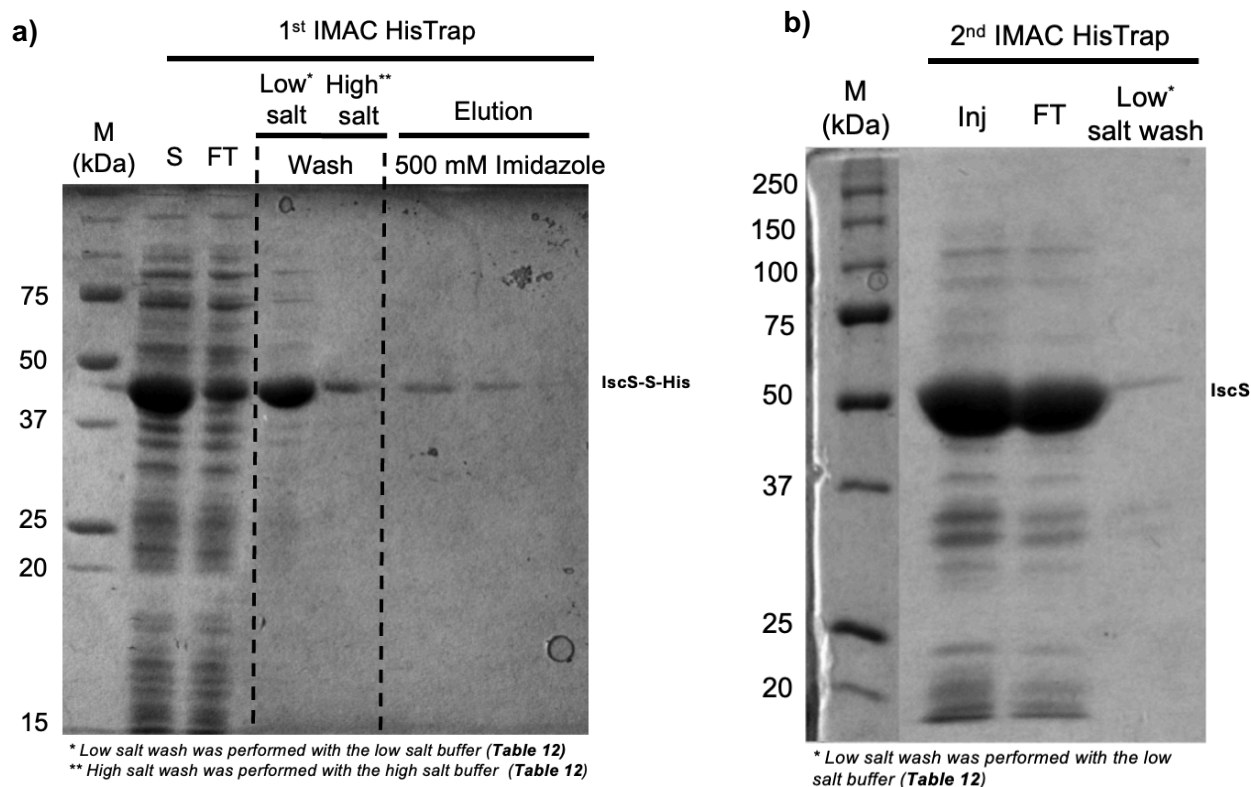


Figure 30. SDS-PAGE gels of IscS after IMAC purification. a) SDS-PAGE analysis of the 1st IMAC purification step of IscS. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad); **S**, filtered supernatant obtained after cell lysis; **FT**, of the loading step of the supernatant onto a Ni²⁺ charged HisTrap™ IMAC HP (Cytiva) 5 mL column; **Low salt wash**; **High salt wash** (Table 12). b) SDS-PAGE analysis of the 2nd IMAC purification step of IscS. The first lane (**M (kDa)**) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad); **Inj**, IscS cleavage mixture after overnight incubation; **FT**, Flow-through obtained after recirculating the IscS cleavage mixture on a Ni²⁺ charged HisTrap column (Cytiva); **Low salt wash** (Table 12).

X-ray crystallography requires high level of sample purity to ensure growth of well-ordered crystals suitable for data collection and structure determination. The presence of contaminants can negatively impact protein crystallization. Therefore, for both quality control and cleaned up of the sample before crystallization, we performed a size exclusion chromatography on a Superdex75 10/300 GL (Cytiva) (**Figure 31**). IscS fractions with the highest degree of purity were pooled and concentrated.

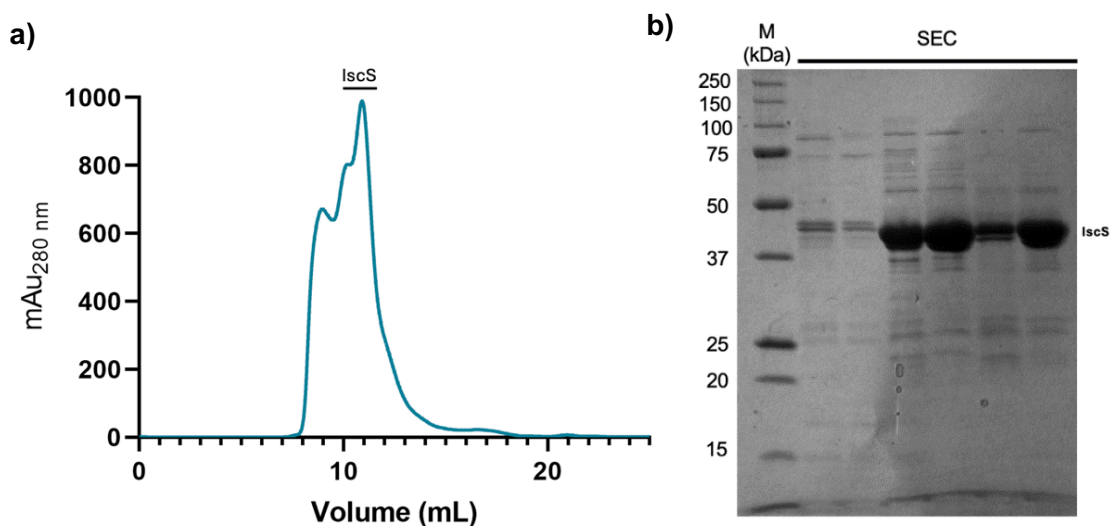


Figure 31. Analysis of IscS purification by SEC. a) Chromatogram of SEC purification of IscS. The blue line represents the absorbance (mAu) at 280 nm. b) SDS-PAGE analysis of the purification of IscS by SEC. The first lane (M (kDa)) corresponds to Precision Plus Protein™ Standards molecular mass marker (Bio-Rad); SEC, Fractions from the SEC purification step.

3.2. IscS crystallization for time-resolved serial crystallography

Time-resolved serial crystallography requires a large number of microcrystals so data can be collected before the highly intense radiation destroys them (“diffraction before destruction” principle). It is therefore necessary to optimize the conditions to produce microcrystal solutions with a density and homogeneity that are suitable for time-resolved serial crystallography experiments.

To identify the optimal condition for microcrystal formation, an initial screening was performed using crystallization solutions from two crystallization kits: MIDAS™ MD1-59 and JCSG+ MD1-37 (Molecular Dimensions). The initial screening was performed using different concentrations of purified protein: 2.49 mg/mL; 5.50 mg/mL; and 6.41 mg/mL. Sitting drop plates for vapor diffusion were used with a protein/precipitant ratio of 1:1 in drops of 2 μ L. Of the crystallization conditions tested, only JCSG+ 1-2 (0.1 M sodium citrate and 20% PEG 3000) from the JCSG+ MD1-37 kit yielded crystals. The crystals appeared after two days (**Figure 32**). However, the initial screening results were not optimal for time-resolved serial crystallography, due to the low crystal density and oversized crystals.

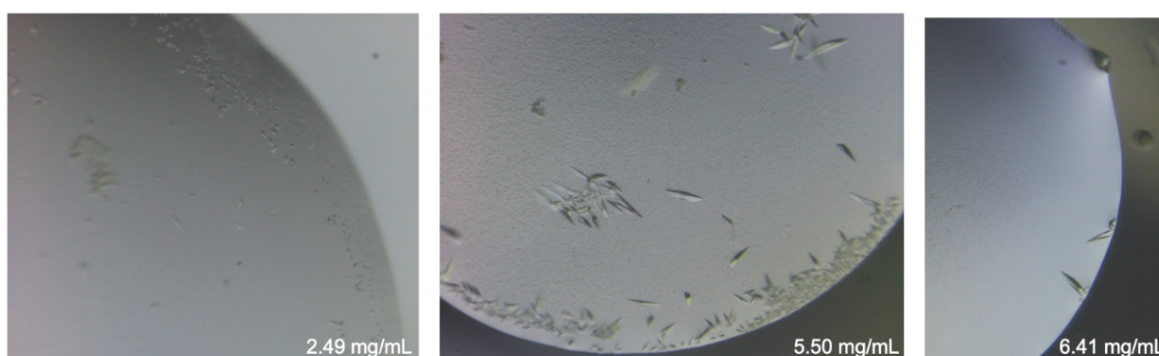


Figure 32. IscS crystals in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions). Crystallization conditions: Different concentrations of purified protein used, 2.40 mg/mL, 5.50 mg/mL and 6.41 mg/mL with 1 mM L-Cysteine. The crystals were produced using the sitting drop vapour diffusion method in a 24 well crystallization plates (Cryschem plates, Hampton Research by mixing 1 μ L of purified protein with 1 μ L of reservoir solution.

To overcome this issue and increase the density of crystals in the drop, the concentration of purified protein was increased to 16.9 mg/mL. This adjustment enhanced nucleation and promoted the production of a larger number of smaller and more homogeneous crystals suitable for time-resolved serial crystallography experiments (**Figure 33**).

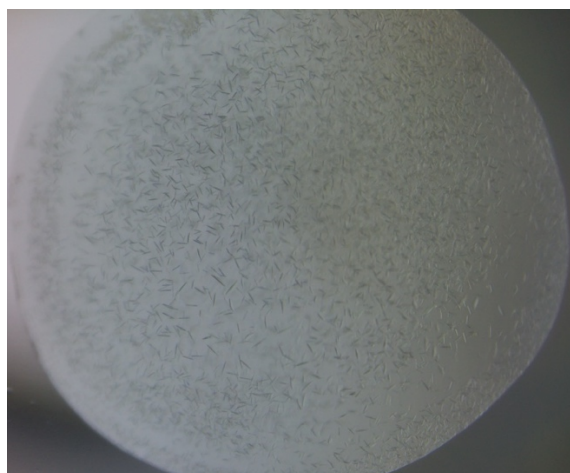


Figure 33. IscS crystals in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions). Crystallization conditions: 16.9 mg/mL plus 1 mM L-Cys. The crystals were produced using the sitting drop vapor diffusion method in a 24-well crystallization plate (Cryschem plates, Hampton Research) by mixing 1 μ L purified protein with 1 μ L reservoir solution.

To standardize the size and shape of the crystals, as well as the density of crystals in the drop, two rounds of seeding were performed. Seeding allows crystals to grow in the metastable region where spontaneous homogeneous nucleation does not occur. This improves control and reproducibility, increasing the likelihood of a successful crystallization experiment. Seeding was performed by crushing the crystals from the previous drops with a Seed Bead™ (Hampton Research) according to the manufacturer's classical protocol to obtain a concentrated seed stock. From this seed stock, several dilutions were made to find the optimal conditions for microcrystal growth (microseeding) (**Figure 34**). For the first round of seeding, the drop contained 1 μ L of purified protein (16.9 mg/mL) and 1 μ L of the seed dilutions. Smaller and more homogeneous crystals were produced from the 10^{-4} and 10^{-6} dilutions (**Figure 34**).

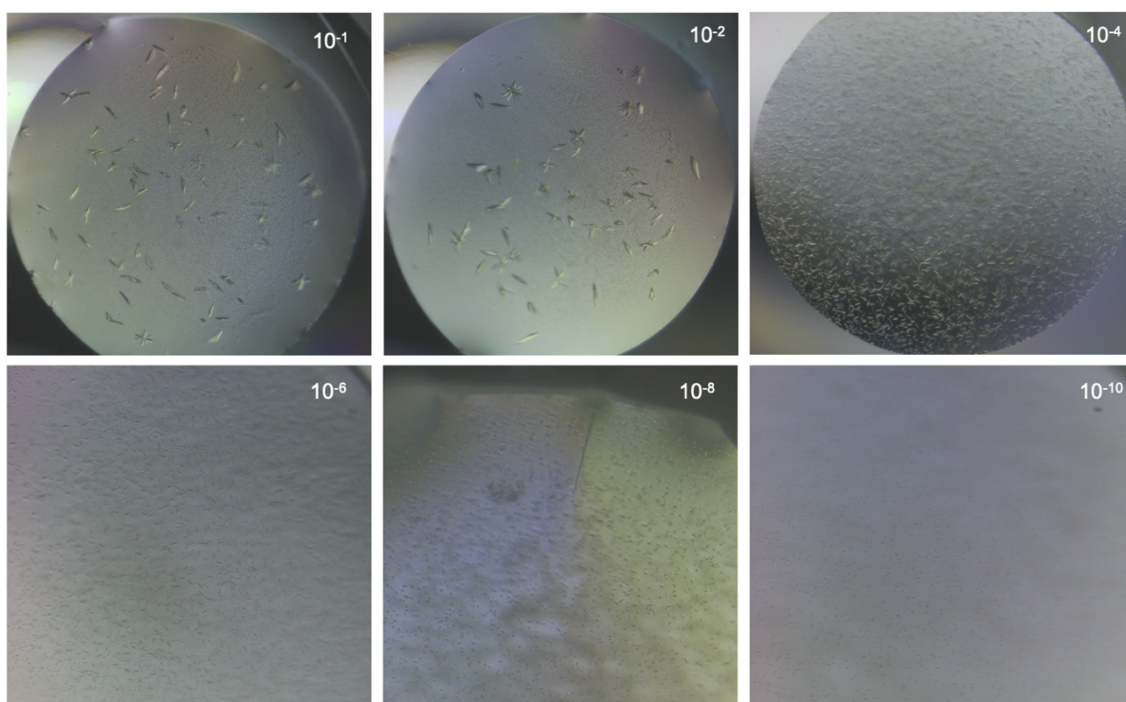


Figure 34. IscS crystals grown in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions) with 1st round seeds. Crystallization conditions: 16.9 mg/mL purified protein plus 1 mM L-Cys and 1st seeds in different dilutions (10^{-1} , 10^{-2} , 10^{-4} , 10^{-6} , 10^{-8} , 10^{-10}). The crystals were produced using the sitting drop vapor diffusion method in a 24-well crystallization plate (Cryschem plates, Hampton Research) by mixing 1 μ L purified protein with 1 μ L 1st round of seeds in different dilutions (10^{-1} , 10^{-2} , 10^{-4} , 10^{-6} , 10^{-8} , 10^{-10}).

The second round of seeding was performed using the crystals grown previously using the 10^{-4} and 10^{-6} seed dilutions to fine tune the size and density of the crystals in the drop. (**Figure 35**). For this round of seeding, the drop contained 3 μL of purified protein (16.9 mg/mL), 2 μL of reservoir solution and 1 μL of seed stock. Adding the reservoir solution in the drop allowed to increase the number of variables that can be finetuned in downstream experiments. The crystals appeared after 24 hours after seeding. For the second-round seeds, the 10^{-6} dilution produced the smallest and most homogenous microcrystals that seem suitable for time-resolved serial crystallography (**Figure 35**).

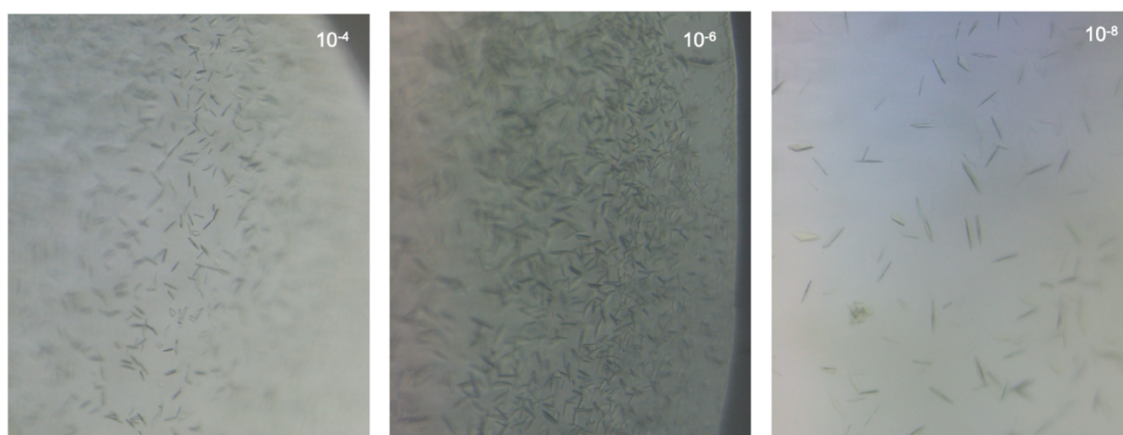


Figure 35. IscS crystals grown in condition 1-2 from JCSG+ MD1-37 (Molecular Dimensions) with 2nd round seeds. Crystallization conditions: 16.9 mg/mL purified protein plus 1 mM L-Cys, 2nd round seeds in different dilutions (10^{-4} , 10^{-6} and 10^{-8}) and reservoir solution. The crystals were produced using the sitting drop vapor diffusion method in a 24-well crystallization plate (Cryschem plates, Hampton Research) by mixing 3 μL purified protein with 2 μL reservoir solution and 1 μL 2nd round seeds in different dilutions (10^{-4} , 10^{-6} and 10^{-8}). Dilution 10^{-6} produced microcrystals suitable for time-resolved serial crystallography.

DISCUSSION

Time-resolved serial crystallography is a major advance in structural biology. It allows molecular processes in proteins to be captured in real time with atomic resolution. However, the success of this technique heavily relies on protocols that enable the collection of high quality diffraction data and guarantee experimental reproducibility.

One key factor to perform reliable and reproducible time-resolved experiments is the consistent production of uniform, high-quality microcrystals. One strategy to produce microcrystals for serial crystallography is microseeding since it yields a high number of microcrystals with minimal size variation (Shoeman et al., 2023).

Using microseeding, we designed a straightforward approach to obtain homogeneous microcrystal preparations that also have suitable density for SFX experiments. This approach involves a two-step seeding process and allows the growth of crystals in the metastable zone. Depending on the concentration of the seeds used, we were able to control the number and the size of the crystals in the crystallization drop. In theory, by increasing the concentration of seeds in the solution, it is possible to increase the nucleation sites and produce homogeneous size microcrystals (Shoeman et al., 2023). Contrary to initial expectations, the size of the IscS crystals obtained was reduced by decreasing the concentration of the seed stock (**Figures 34 and 35**). In order to confirm this finding and ensure reproducibility, it is necessary to repeat the experiment.

The greatest advantage of seeding is that can be applied to any other protein to produce large quantities of homogeneously sized microcrystals. This technique was also applied to trypsin in the lab, and after two-rounds of seeding we were able to successfully obtain homogeneous microcrystals (**data not shown**). In this case, the smallest microcrystals were obtained when using higher concentrations of the seed stock, as expected.

Other important factor that plays a critical role in the production of microcrystals is the concentration of purified protein. Increasing the concentration of purified protein in the solution should lead to an increase in nucleation sites, which facilitates the production of a higher amount of microcrystals. By increasing the concentration of protein, we were able to produce more and smaller crystals as shown in **Figure 33**. This aspect becomes particularly important when transitioning from the vapor diffusion method to the batch method for larger-scale production. In the batch method, it is important to achieve the supersaturation state as quickly as possible to rapidly promote protein nucleation.

Taken together, our results show microseeding as a straightforward approach to produce homogeneous microcrystals suitable for serial crystallography. After several steps of seeding, it was possible to obtain homogeneous microcrystals of IscS and trypsin.

Moving forward it will be necessary to guarantee the reproducibility of the experimental protocols set up to produce microcrystals suitable for time resolved serial-crystallography, including the ones being set for trypsin and later for thrombin. Every protein will crystallize under specific conditions that will have to be previously screened and optimized, but having a seeding protocol that is likely to yield homogeneous preparations of microcrystals is a new and auspicious starting point for future experiments taking place at the ID29 beamline (ESRF-EBS, France).

Conclusions and Future Perspectives

To fully understand the function of macromolecules, it is essential to combine structural information with an understating of the dynamic changes that occur over time in the protein. Time-resolved serial crystallography can expand our knowledge of protein function by making these changes visible. This technique allows the detailed characterization of intermediate states and reaction steps by the determination of both the chemical and kinetics mechanism.

In serial crystallography, the diffraction data is collected in a serial fashion using extremely bright and fast X-ray sources before the radiation has time to destroy the sample ("diffraction before destruction" principle). The data are collected from very small and homogeneous crystals (microcrystals or nanocrystals). The effectiveness of time-resolved crystallography depends on optimizing protocols to ensure data quality, improve experimental efficiency, as well the ability to capture dynamic processes. Because even small variations can affect the structural data obtained, it is critical to ensure experimental reproducibility and standardization.

In this work, we present a straightforward strategy to produce large quantities of homogeneous microcrystals suitable for SSX, since one of the important factors in performing reliable and reproducible experiments is the use of homogeneous, high-quality microcrystals. Using successive rounds of seeding, we were able to transition from large crystals to microcrystals. However, the size of the IscS crystals was reduced by using more diluted seed solutions with the vapor diffusion method, contrary to our initial expectations and it is therefore necessary to repeat the experiments to ensure reproducibility. The same approach has yielded homogeneous microcrystals of trypsin, showing this is a robust method that can be potentially used in future SSX experiments.

Since serial crystallography requires a large amount of crystals for data collection, we plan to scale up the batch method for both IscS and trypsin proteins in order to obtain a larger volume of highly homogeneous microcrystals. In addition, we plan to optimize time-resolved experiments using microcrystals of these proteins in complex with photoactivatable substrates at the ID29 beamline of ESRF-EBS (Grenoble, France), with the ultimate goal of applying these protocols to study the proteolytic mechanism of human thrombin and establish their applicability to study a larger set of enzymes.

The methodologies established with this work will not only provide a real-world benchmark for the performance of the recently established ID29 beamline, but will also

define an experimental approach for future studies using time-resolved serial crystallography at room temperature in synchrotron light sources

References

- Abro, R., Kiran, N., Ahmed, S., Muhammad, A., Jatoi, A. S., Mazari, S. A., Salma, U., & Plechkova, N. V. (2022). Extractive desulfurization of fuel oils using deep eutectic solvents – A comprehensive review. *Journal of Environmental Chemical Engineering*, 10(3), 107369. <https://doi.org/10.1016/j.jece.2022.107369>
- Aloh, C. H., Zeczycki, T. N., & Ellis, H. R. (2023). Oligomeric Changes Regulate Flavin Transfer in Two-Component FMN Reductases Involved in Sulfur Metabolism. *Biochemistry*, 62(18), 2751–2762. <https://doi.org/10.1021/acs.biochem.3c00361>
- Beale, J. H., Bolton, R., Marshall, S. A., Beale, E. V., Carr, S. B., Ebrahim, A., Moreno-Chicano, T., Hough, M. A., Worrall, J. A. R., Tews, I., & Owen, R. L. (2019). Successful sample preparation for serial crystallography experiments. *Journal of Applied Crystallography*, 52(6), 1385–1396. <https://doi.org/10.1107/S1600576719013517>
- Boniek, D., Figueiredo, D., Dos Santos, A. F. B., & De Resende Stoianoff, M. A. (2015). Biodesulfurization: A mini review about the immediate search for the future technology. *Clean Technologies and Environmental Policy*, 17(1), 29–37. <https://doi.org/10.1007/s10098-014-0812-x>
- Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I., & Doak, R. B. (2015). Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. *Acta Crystallographica Section D Biological Crystallography*, 71(2), 387–397. <https://doi.org/10.1107/S1399004714026327>
- Brito, J. A., & Archer, M. (2020). Structural biology techniques: X-ray crystallography, cryo-electron microscopy, and small-angle X-ray scattering. In *Practical Approaches to Biological Inorganic Chemistry* (pp. 375–416). Elsevier. <https://doi.org/10.1016/B978-0-444-64225-7.00010-9>
- Chayen, N. E., & Saridakis, E. (2008). Protein crystallization: From purified protein to diffraction-quality crystal. *Nature Methods*, 5(2), 147–153. <https://doi.org/10.1038/nmeth.f.203>

Chen, S., Zhao, C., Liu, Q., Zang, M., Liu, C., & Zhang, Y. (2018). Thermophilic biodesulfurization and its application in oil desulfurization. *Applied Microbiology and Biotechnology*, 102(21), 9089–9103. <https://doi.org/10.1007/s00253-018-9342-5>

Cowtan, K. (2003). Phase Problem in X-ray Crystallography, and Its Solution. In Wiley, *Encyclopedia of Life Sciences* (1.^a ed.). Wiley. <https://doi.org/10.1038/npg.els.0002722>

Deller, M. C., Kong, L., & Rupp, B. (2016). Protein stability: A crystallographer's perspective. *Acta Crystallographica Section F Structural Biology Communications*, 72(2), 72–95. <https://doi.org/10.1107/S2053230X15024619>

Demirbas, A., & Bamufleh, H. S. (2017). Optimization of crude oil refining products to valuable fuel blends. *Petroleum Science and Technology*, 35(4), 406–412. <https://doi.org/10.1080/10916466.2016.1261162>

Ferreira, P., Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2017). Improving the Catalytic Power of the DszD Enzyme for the Biodesulfurization of Crude Oil and Derivatives. *Chemistry – A European Journal*, 23(68), 17231–17241. <https://doi.org/10.1002/chem.201704057>

Gangwal, S. K. (2011). Desulfurization for Fuel Cells. In *Fuel Cells: Technologies for Fuel Processing* (pp. 317–360). Elsevier. <https://doi.org/10.1016/B978-0-444-53563-4.10011-2>

Gräslund, S., Sagemark, J., Berglund, H., Dahlgren, L.-G., Flores, A., Hammarström, M., Johansson, I., Kotenyova, T., Nilsson, M., Nordlund, P., & Weigelt, J. (2008). The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. *Protein Expression and Purification*, 58(2), 210–221. <https://doi.org/10.1016/j.pep.2007.11.008>

Gray, K. A., Mrachko, G. T., & Squires, C. H. (2003). Biodesulfurization of fossil fuels. *Current Opinion in Microbiology*, 6(3), 229–235. [https://doi.org/10.1016/S1369-5274\(03\)00065-1](https://doi.org/10.1016/S1369-5274(03)00065-1)

Gray, K. A., Pogrebinsky, O. S., Mrachko, G. T., Xi, L., Monticello, D. J., & Squires, C. H. (1996). Molecular mechanisms of biocatalytic desulfurization of fossil fuels. *Nature*

Biotechnology, 14(13), 1705–1709. <https://doi.org/10.1038/nbt1296-1705>

Grossman, M. J., Lee, M. K., Prince, R. C., Garrett, K. K., George, G. N., & Pickering, I. J. (1999). Microbial Desulfurization of a Crude Oil Middle-Distillate Fraction: Analysis of the Extent of Sulfur Removal and the Effect of Removal on Remaining Sulfur. *Applied and Environmental Microbiology*, 65(1), 181–188. <https://doi.org/10.1128/AEM.65.1.181-188.1999>

Grünbein, M. L., & Nass Kovacs, G. (2019). Sample delivery for serial crystallography at free-electron lasers and synchrotrons. *Acta Crystallographica Section D Structural Biology*, 75(2), 178–191. <https://doi.org/10.1107/S205979831801567X>

Hadian-Jazi, M., Berntsen, P., Marman, H., Abbey, B., & Darmanin, C. (2021). Analysis of Multi-Hit Crystals in Serial Synchrotron Crystallography Experiments Using High-Viscosity Injectors. *Crystals*, 11(1), 49. <https://doi.org/10.3390/cryst11010049>

Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., & Shewchuk, L. M. (2007). Crystallization of protein–ligand complexes. *Acta Crystallographica Section D Biological Crystallography*, 63(1), 72–79. <https://doi.org/10.1107/S0907444906047020>

Held, Paul. (2021, September 1). *Determination of NADH Concentrations with the Synergy™ 2 Multi-Detection Microplate Reader using Fluorescence or Absorbance*. Retrieved from <https://www.agilent.com/cs/library/applications/determining-NADH-concentrations-5994-2652EN-agilent.pdf>

Ilari, A., & Savino, C. (2008). Protein Structure Determination by X-Ray Crystallography. In J. M. Keith (Ed.), *Bioinformatics* (Vol. 452, pp. 63–87). Humana Press. https://doi.org/10.1007/978-1-60327-159-2_3

Javadli, R., & De Klerk, A. (2012). Desulfurization of heavy oil. *Applied Petrochemical Research*, 1(1–4), 3–19. <https://doi.org/10.1007/s13203-012-0006-6>

Jha, D., Maheshwari, P., Singh, Y., Haider, M. B., Kumar, R., & Balathanigaimani, M. S. (2023). A comparative review of extractive desulfurization using designer solvents: Ionic liquids & deep eutectic solvents. *Journal of the Energy Institute*, 110, 101313.

<https://doi.org/10.1016/j.joei.2023.101313>

Kalita, M., Sangannavar, P. A., Chutia, M., Jha, D. K., Sathyanarayana, K., Kumar, J. S., & Subrahmanyam, G. (2022). Microbial biodesulfurization. In *Microbial Resource Technologies for Sustainable Development* (pp. 333–351). Elsevier. <https://doi.org/10.1016/B978-0-323-90590-9.00019-5>

Kamali, N., Tavallaie, M., Bambai, B., Karkhane, A. A., & Miri, M. (2010). Site-directed mutagenesis enhances the activity of NADH-FMN oxidoreductase (DszD) activity of *Rhodococcus erythropolis*. *Biotechnology Letters*, 32(7), 921–927. <https://doi.org/10.1007/s10529-010-0254-4>

Kim, B. H., Kim, H. Y., Kim, T. S., & Park, D. H. (1995). Selectivity of desulfurization activity of *Desulfovibrio desulfuricans* M6 on different petroleum products. *Fuel Processing Technology*, 43(1), 87–94. [https://doi.org/10.1016/0378-3820\(95\)00010-5](https://doi.org/10.1016/0378-3820(95)00010-5)

Krissinel, E., & Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, 372(3), 774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>

Lee, K. X., & Valla, J. A. (2019). Adsorptive desulfurization of liquid hydrocarbons using zeolite-based sorbents: A comprehensive review. *Reaction Chemistry & Engineering*, 4(8), 1357–1386. <https://doi.org/10.1039/C9RE00036D>

Li, F.-L., Su, W. B., Tao, Q.-L., Zhang, L.-Y., & Zhang, Y.-W. (2020). Expression, biochemical characterization, and mutation of a water forming NADH: FMN oxidoreductase from *Lactobacillus rhamnosus*. *Enzyme and Microbial Technology*, 134, 109464. <https://doi.org/10.1016/j.enzmictec.2019.109464>

Lin, C.-W., McCabe, J. W., Russell, D. H., & Barondeau, D. P. (2020). Molecular Mechanism of ISC Iron–Sulfur Cluster Biogenesis Revealed by High-Resolution Native Mass Spectrometry. *Journal of the American Chemical Society*, 142(13), 6018–6029. <https://doi.org/10.1021/jacs.9b11454>

Liu, H., Zhao, Y., & Sun, J. (2023). Heterogeneous Nucleation in Protein Crystallization. *Biomimetics*, 8(1), 68. <https://doi.org/10.3390/biomimetics8010068>

Ma, T. (2010). The Desulfurization Pathway in *Rhodococcus*. In H. M. Alvarez (Ed.), *Biology of Rhodococcus* (Vol. 16, pp. 207–230). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12937-7_8

Majeed, S., Ofek, G., Belachew, A., Huang, C., Zhou, T., & Kwong, P. D. (2003). Enhancing Protein Crystallization through Precipitant Synergy. *Structure*, 11(9), 1061–1070. [https://doi.org/10.1016/S0969-2126\(03\)00185-0](https://doi.org/10.1016/S0969-2126(03)00185-0)

Mamuad, R. Y., & Choi, A. E. S. (2023). Biodesulfurization Processes for the Removal of Sulfur from Diesel Oil: A Perspective Report. *Energies*, 16(6), 2738. <https://doi.org/10.3390/en16062738>

Martínez, I., García, J. L., & Díaz, E. (2017). Genetic Engineering for Removal of Sulfur from Fuel Aromatic Heterocycles. Em S. Y. Lee (Ed.), *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals* (pp. 183–204). Springer International Publishing. https://doi.org/10.1007/978-3-319-50436-0_206

Mcperson, A. (2004). Introduction to protein crystallization. *Methods*, 34(3), 254–265. <https://doi.org/10.1016/j.ymeth.2004.03.019>

Meszaros, P., & Westenhoff, S. (2024). Time-resolved serial crystallography to reveal protein structural changes. *Trends in Biochemical Sciences*, 49(2), 183–184. <https://doi.org/10.1016/j.tibs.2023.09.009>

Moghadam, F. R., Kianpour, E., Azizian, S., Yarie, M., & Zolfigol, M. A. (2020). Extractive desulfurization of liquid fuel using diamine-terminated polyethylene glycol as a very low vapour pressure and green molecular solvent. *Royal Society Open Science*, 7(11), 200803. <https://doi.org/10.1098/rsos.200803>

Mohebbi, G., & Ball, A. S. (2008). Biocatalytic desulfurization (BDS) of petrodiesel fuels. *Microbiology*, 154(8), 2169–2183. <https://doi.org/10.1099/mic.0.2008/017608-0>

Mohebbi, G., & Ball, A. S. (2016). Biodesulfurization of diesel fuels – Past, present and future perspectives. *International Biodeterioration & Biodegradation*, 110, 163–180.

<https://doi.org/10.1016/j.ibiod.2016.03.011>

Monteiro, D. C. F., Amoah, E., Rogers, C., & Pearson, A. R. (2021). Using photocaging for fast time-resolved structural biology studies. *Acta Crystallographica Section D Structural Biology*, 77(10), 1218–1232. <https://doi.org/10.1107/S2059798321008809>

Monticello, D. J. (2000). Biodesulfurization and the upgrading of petroleum distillates. *Current Opinion in Biotechnology*, 11(6), 540–546. [https://doi.org/10.1016/S0958-1669\(00\)00154-3](https://doi.org/10.1016/S0958-1669(00)00154-3)

Morris, R., Black, K. A., & Stollar, E. J. (2022). Uncovering protein function: From classification to complexes. *Essays in Biochemistry*, 66(3), 255–285. <https://doi.org/10.1042/EBC20200108>

Nam, K. H. (2024). Guide to serial synchrotron crystallography. *Current Research in Structural Biology*, 7, 100131. <https://doi.org/10.1016/j.crstbi.2024.100131>

Nieto-Taype, M. A., Garcia-Ortega, X., Albiol, J., Montesinos-Seguí, J. L., & Valero, F. (2020). Continuous Cultivation as a Tool Toward the Rational Bioprocess Development With *Pichia Pastoris* Cell Factory. *Frontiers in Bioengineering and Biotechnology*, 8, 632. <https://doi.org/10.3389/fbioe.2020.00632>

Ohshiro, T., & Izumi, Y. (1999). Microbial Desulfurization of Organic Sulfur Compounds in Petroleum. *Bioscience, Biotechnology, and Biochemistry*, 63(1), 1–9. <https://doi.org/10.1271/bbb.63.1>

Omar, R. A., & Verma, N. (2022). Review of Adsorptive Desulfurization of Liquid Fuels and Regeneration Attempts. *Industrial & Engineering Chemistry Research*, 61(25), 8595–8606. <https://doi.org/10.1021/acs.iecr.2c01426>

Pandey, S., Poudyal, I., & Malla, T. N. (2020). Pump-Probe Time-Resolved Serial Femtosecond Crystallography at X-Ray Free Electron Lasers. *Crystals*, 10(7), 628. <https://doi.org/10.3390/cryst10070628>

Pearson, A. R., & Mehrabi, P. (2020). Serial synchrotron crystallography for time-resolved structural biology. *Current Opinion in Structural Biology*, 65, 168–174.

<https://doi.org/10.1016/j.sbi.2020.06.019>

Poddar, H., Heyes, D. J., Schirò, G., Weik, M., Leys, D., & Scrutton, N. S. (2022). A guide to time-resolved structural analysis of light-activated proteins. *The FEBS Journal*, 289(3), 576–595. <https://doi.org/10.1111/febs.15880>

Pusey, M. L., & Aygün, R. S. (2017). *Data Analytics for Protein Crystallization* (Vol. 25). Springer International Publishing. <https://doi.org/10.1007/978-3-319-58937-4>

Rajendran, A., Cui, T., Fan, H., Yang, Z., Feng, J., & Li, W. (2020). A comprehensive review on oxidative desulfurization catalysts targeting clean energy and environment. *Journal of Materials Chemistry A*, 8(5), 2246–2285. <https://doi.org/10.1039/C9TA12555H>

Rondeau, J.-M., & Schreuder, H. (2015). Protein Crystallography and Drug Discovery. In *The Practice of Medicinal Chemistry* (pp. 511–537). Elsevier. <https://doi.org/10.1016/B978-0-12-417205-0.00022-5>

Russo Krauss, I., Merlino, A., Vergara, A., & Sica, F. (2013). An Overview of Biological Macromolecule Crystallization. *International Journal of Molecular Sciences*, 14(6), 11643–11691. <https://doi.org/10.3390/ijms140611643>

Sadare, O., Obazu, F., & Daramola, M. (2017). Biodesulfurization of Petroleum Distillates—Current Status, Opportunities and Future Challenges. *Environments*, 4(4), 85. <https://doi.org/10.3390/environments4040085>

Saha, B., Vedachalam, S., & Dalai, A. K. (2021). Review on recent advances in adsorptive desulfurization. *Fuel Processing Technology*, 214, 106685. <https://doi.org/10.1016/j.fuproc.2020.106685>

Schulz, E. C., Yorke, B. A., Pearson, A. R., & Mehrabi, P. (2022). Best practices for time-resolved serial synchrotron crystallography. *Acta Crystallographica Section D Structural Biology*, 78(1), 14–29. <https://doi.org/10.1107/S2059798321011621>

Sharma, S., Ebrahim, A., & Keedy, D. A. (2023). Room-temperature serial synchrotron

crystallography of the human phosphatase PTP1B. *Acta Crystallographica Section F Structural Biology Communications*, 79(1), 23–30. <https://doi.org/10.1107/S2053230X22011645>

Shi, Q., & Wu, J. (2021). Review on Sulfur Compounds in Petroleum and Its Products: State-of-the-Art and Perspectives. *Energy & Fuels*, 35(18), 14445–14461. <https://doi.org/10.1021/acs.energyfuels.1c02229>

Shoeman, R. L., Hartmann, E., & Schlichting, I. (2023). Growing and making nano- and microcrystals. *Nature Protocols*, 18(3), 854–882. <https://doi.org/10.1038/s41596-022-00777-5>

Soleimani, M., Bassi, A., & Margaritis, A. (2007). Biodesulfurization of refractory organic sulfur compounds in fossil fuels. *Biotechnology Advances*, 25(6), 570–596. <https://doi.org/10.1016/j.biotechadv.2007.07.003>

Sousa, S. F., Sousa, J. F. M., Barbosa, A. C. C., Ferreira, C. E., Neves, R. P. P., Ribeiro, A. J. M., Fernandes, P. A., & Ramos, M. J. (2016). Improving the Biodesulfurization of Crude Oil and Derivatives: A QM/MM Investigation of the Catalytic Mechanism of NADH-FMN Oxidoreductase (DszD). *The Journal of Physical Chemistry A*, 120(27), 5300–5306. <https://doi.org/10.1021/acs.jpca.6b01536>

Sprenger, J., Carey, J., Schulz, A., Drouard, F., Lawson, C. L., Von Wachenfeldt, C., Linse, S., & Lo Leggio, L. (2021). Guest-protein incorporation into solvent channels of a protein host crystal (hostal). *Acta Crystallographica Section D Structural Biology*, 77(4), 471–485. <https://doi.org/10.1107/S2059798321001078>

Šrajer, V., & Schmidt, M. (2017). Watching proteins function with time-resolved x-ray crystallography. *Journal of Physics D: Applied Physics*, 50(37), 373001. <https://doi.org/10.1088/1361-6463/aa7d32>

Stollar, E. J., & Smith, D. P. (2020). Uncovering protein structure. *Essays in Biochemistry*, 64(4), 649–680. <https://doi.org/10.1042/EBC20190042>

Stylianou, M., Vyrides, I., & Agapiou, A. (2021). Oil biodesulfurization: A review of applied analytical techniques. *Journal of Chromatography B*, 1171, 122602.

<https://doi.org/10.1016/j.jchromb.2021.122602>

Wang, H., & Wang, J. (2017). How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Science*, 26(1), 32–39. <https://doi.org/10.1002/pro.3022>

Wang, L., Ji, G., & Huang, S. (2019). Contribution of the Kodama and 4S pathways to the dibenzothiophene biodegradation in different coastal wetlands under different C/N ratios. *Journal of Environmental Sciences*, 76, 217–226. <https://doi.org/10.1016/j.jes.2018.04.029>

Wang, Y., Chu, J., Zhuang, Y., Wang, Y., Xia, J., & Zhang, S. (2009). Industrial bioprocess control and optimization in the context of systems biotechnology. *Biotechnology Advances*, 27(6), 989–995. <https://doi.org/10.1016/j.biotechadv.2009.05.022>

Annexes

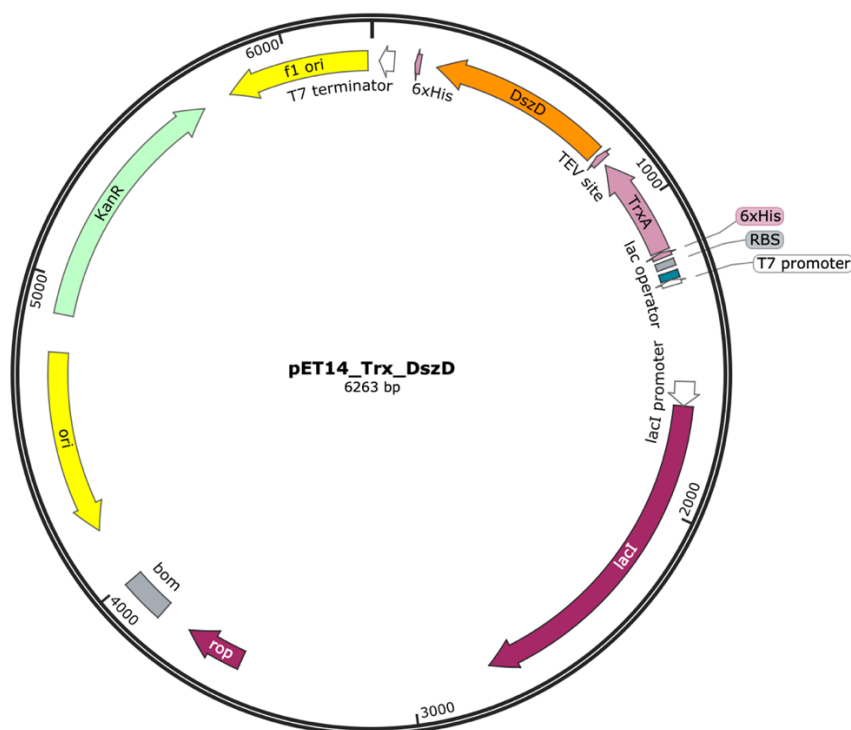


Figure 36. Map of the plasmid used for recombination expression of DszD *wt* and variants (T62D, T62R, T62L, and T62Y). The pET14_Trx_DszD contains a KAN resistance gene and the expression of DszD is under the control of the T7 promoter. DszD is expressed as a fusion protein with a Trx as solubility tag and histidine tag as an affinity tag.

Table 13. Composition of the crystallization conditions used from Morpheus® MD1-46 (Molecular Dimensions) conditions. Composition of the mix of additives, buffer systems and precipitant mixes used from Morpheus® MD1-46 (Molecular Dimensions) conditions.

Mix name		Composition
Mix of additives	NPS*	0.3 M Sodium nitrate, 0.3 M Sodium phosphate dibasic, 0.3 M Ammonium sulfate
	Carboxylic acids	0.2 M Sodium formate; 0.2 M Ammonium acetate; 0.2 M Sodium citrate tribasic dihydrate; 0.2 M Potassium sodium tartrate tetrahydrate; 0.2 M Sodium oxamate
Buffer Systems	Buffer system 1	1.0 M (Imidazole; MES; monohydrate (acid)) pH 6.5
Precipitant mix	Precipitant mix 3	40% (v/v) Glycerol; 20% (w/v) PEG 4000

*Nitrate phosphate sulfate