



**UNIVERSITY  
OF TURKU**

# **The iron-entry pores in Dps-like proteins as a potential drug-design target: Structural investigations**

Faculty of Medicine,  
Institute of Biomedicine  
Master's thesis

Author:  
Polychronis Tatsis

Supervisor:  
Group Leader, Senior Research Fellow, Adjunct Professor, Tassos Papageorgiou

30.05.2022

Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

**This thesis is dedicated to my grandfather, Polychronis.**

For his endless support between our uncountable afternoon Greek coffees.

Master's thesis

**Subject:** Crystallography, Structural Biology

**Author(s):** Polychronis Tatsis

**Title:** The iron-entry pores in Dps-like proteins as a potential drug-design target: Structural investigations

**Supervisor(s):** Group Leader, Senior Research Fellow, Adjunct Professor, Tassos Papageorgiou

**Number of pages:** 53 pages

**Date:** 30.05.2022

### Abstract

*Streptococcus suis* is an emerging catalase-negative zoonotic pathogen that uses the peroxide resistance protein (Dpr) as a defensive mechanism against oxidative stress. Dpr belongs to a family of proteins that form spherical dodecamers with a hollow cavity in the middle. Dpr, as other members of the family, uses four pores found on the surface of the dodecamer and formed by the N-terminals of adjacent monomers (N-terminal pores) to take up  $\text{Fe}^{2+}$  and deposit it inside the cavity after its oxidation to  $\text{Fe}^{3+}$  in ferroxidase sites in the interior of the dodecamer. In this way, the generation of toxic hydroxyl radicals via Fenton's reaction is avoided. In this study, a new purification process and crystallization conditions for Dpr were found. Besides, the ligandability of Dpr for use as a drug target was investigated. 6xHis-tagged Dpr was successfully produced and purified. Crystallization screens yielded crystals in 10 conditions and further optimization led to crystals suitable for structural analysis. Synchrotron X-ray data were collected to 2.2 Å resolution. A novel ligand library design led to an initial library of 82 compounds that could act as possible N-terminal pore blockers. After a score threshold of -7, twenty (20) ligands remained. Similar, to the latter ones, marketed ligands were retrieved, and ten (10) of them were kept, all sharing the feature of having aromatic rings. Phe133 was found as the only residue responsible for Pi-pi interactions with the ligands. This is the first successful approach for 6xHis-tag Dpr crystal production and structure determination. It is also the first approach for ligand creation against the N-terminal pores of the Dpr, setting the basis for new possible future therapeutic approaches for *S. suis*-related infections treatment, avoiding the obstacle of antibiotic resistance.

**Keywords:** oxidative stress, DNA-binding protein, antibiotic resistance, drug design, *Streptococcus suis*, meningitis, ferroxidase, crystallization.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Iron & Oxidative stress	7
1.2	Dps & Dps-like proteins	8
1.3	Structural Characterization	9
1.4	Iron Incorporation	10
1.4.1	Iron Entry Mechanism	10
1.4.2	Ferroxidase site	11
1.4.3	Iron core formation	12
1.5	DNA Binding	12
1.6	Applications of the Iron Core	13
1.7	Bacterial Infections & Antibiotic Resistance	14
<b>2</b>	<b>Aims</b>	<b>16</b>
<b>3</b>	<b>Materials &amp; Methods</b>	<b>17</b>
3.1	Protein Crystallization & Structure Determination	17
3.1.1	Bacterial Culture & Protein Purification	17
3.1.2	Protein Purification & Concentration Determination	17
3.1.3	Batch One	17
3.1.4	Batch Two	18
3.1.5	Storage	19
3.1.6	Crystallization Screenings	19
3.1.7	Scale-up & crystallization optimization	20
3.1.8	IceBear	20
3.1.9	Stereoscope Visualization & Crystal Collection	20
3.1.10	X-ray Diffraction	21
3.2	Computational/Ligandability Part	21
<b>4</b>	<b>Results</b>	<b>22</b>
4.1	Protein Purification, Concentration Determination & Storage	22
4.2	Protein Crystallization & Structure Determination	24
4.2.1	First Batch	24
4.2.2	Second Batch	27
4.2.3	X-ray data collection	29
4.3	Computational/Ligandability Part	30

4.3.1	Ligand Library Creation .....	31
<b>5</b>	<b>Discussion .....</b>	<b>41</b>
<b>6</b>	<b>Conclusions .....</b>	<b>43</b>
<b>7</b>	<b>Abbreviations .....</b>	<b>44</b>
	<b>References .....</b>	<b>45</b>
<b>8</b>	<b>Appendices.....</b>	<b>51</b>
8.1	Appendix 1. Master's Thesis Supervision Plan.....	51
8.2	Appendix 2. Mandatory mid-phase feedback.....	53



# 1 Introduction

## 1.1 Iron & Oxidative stress

Iron is the most abundant chemical element on Earth by mass, comprising about 80% of the inner and outer cores of Earth. In addition, it is one of the most important micronutrients, playing a vital role in human health. It is included in numerous dietary products and its deficiency can lead to life-long health-related issues, with iron deficiency being one of the most characteristic (Frey & Reed, 2012; Miller, 2013).

From a different point of view, iron availability, on the contrary, is considered to be a growth-limiting factor for bacteria. Iron is widely distributed in cells and tends to get oxidized from  $\text{Fe}^{2+}$  to insoluble  $\text{Fe}^{3+}$ . Moreover, this oxidization results in the disruption of normal cellular functions. Therefore, most of the iron in the cytosol is attached to different proteins to form various complexes. The final concentration of free  $\text{Fe}^{3+}$  and  $\text{Fe}^{2+}$  is as low as  $10^{-18}$  and  $10^{-8}$  M, respectively (R. J. P. Williams, 1982).

In addition, the presence of  $\text{Fe}^{2+}$  can be the starting point for the production of toxic hydroxyl radicals. These free radicals are reactive oxygen species (ROS) and are associated with numerous actions such as induction of DNA breakage, lipid peroxidation, and finally degradation of the biomolecules. Generation of the abovementioned radicals can occur based on Fenton's reaction (Imlay, 2008; Storz & Imlay, 1999):



Taking into consideration all of the above, it is understood that oxidative stress can cause bacterial cell damage. Accordingly, bacteria have developed various mechanisms to protect themselves from oxidative stress (Lushchak, 2001). These mechanisms are mainly divided into three categories: a) prevention of ROS generation; b) quenching of chain propagation (regeneration of a reactive intermediate during the course of a chain reaction) c) repair of damage. More specifically, non-enzymatic small molecules, such as glutathione (GSH)—a Glu-Cys-Gly tripeptide—and thioredoxin, comprise the initial line of defense of the bacteria for protection from ROS. In addition, mechanisms containing superoxide dismutase, catalases, and alkyl hydroperoxide reductase are also utilized for ROS protection (Lemire et al., 2017).

## 1.2 Dps & Dps-like proteins

Many bacteria, including *Streptococcus suis*, lack catalase, the primary enzyme for peroxidase resistance, and alternatively, they use another mechanism for their protection from ROS (Karas et al., 2015; Kauko et al., 2006). This mechanism relies on the so-called DNA-binding protein from starved cells (Dps) proteins. These proteins are within a well-conserved family of proteins, which are widely distributed in the bacterial and archaea kingdom (Franceschini et al., 2006; Haikarainen & Papageorgiou, 2009).

Dps proteins belong to the ferritin protein family due to the fact that they share many common conserved regions. Dps proteins are also known as mini-ferritins owing to their smaller size compared to ferritins. Indeed, ferritins have a total of 24 subunits, while Dps-like proteins consist of 12 subunits. As a result, Dps proteins create one of the smallest spherical protein cages known so far (Aumiller et al., 2018; Grant et al., 1998).

Regarding the protecting mechanisms of Dps proteins, two modes have been described: i) binding to  $\text{Fe}^{2+}$  ions and consequently preventing the formation of toxic hydroxyl radicals by Fenton's reaction and/or ii) binding to the DNA, thus creating a protective layer against harmful oxidative radicals. Depending on where these Dps proteins derive from, some of them have a dual action, while others store the iron in a bioavailable form without binding to the DNA; for example, *Listeria innocua* Dps does not bind to DNA while *E. coli* Dps does (Bozzi et al., 1997; Haikarainen & Papageorgiou, 2009).

Dps proteins are a larger family of proteins that include a subdivision of proteins called Dps-like proteins. Just as their name explains, these proteins have a similar mode of action as the Dps protein and they are crucial factors for the protection of prokaryotic cells from oxidative damage. Currently, more than a thousand Dps-like proteins have been identified, with an approximate distribution of 97% found in bacteria and 3% in archaea. Besides the already mentioned two advantages of DNA-binding and the iron core formation, these proteins exhibit also additional functions, including cell protection against a variety of stresses, such as metal stress, cold and heat shock, and nutritional deprivation. Furthermore, they have been implicated in inflammatory processes and play a vital role for future use in biotechnology and nanotechnology, opening a new spectrum of applications for the pharma industry (Hébraud &

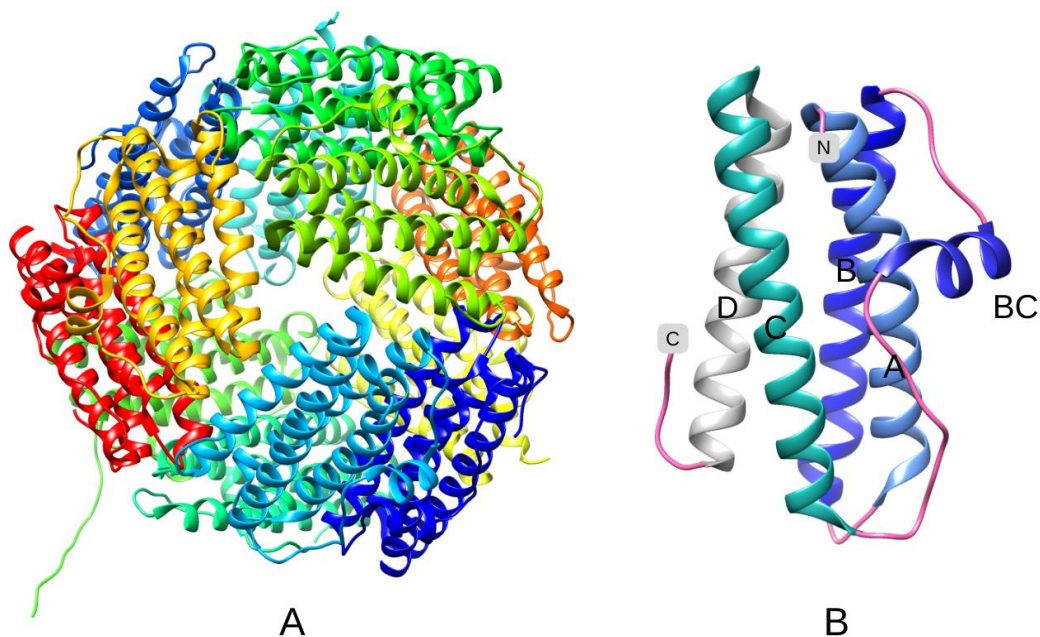


Guzzo, 2000; Nicodème et al., 2004; Perrin et al., 1999; Surgutskaya et al., 2017; Y. N. Wu et al., 2018).

### 1.3 Structural Characterization

Dps-like proteins form a dodecameric quaternary structure as depicted in Figure 1. Each subunit of the dodecamer has a molecular weight (MW) of about 20 kDa and forms a four- $\alpha$ -helix bundle. This folding is similar to that of ferritins, a fact that supports the theory that Dps-like proteins and ferritins share a common evolutionary ancestor (Andrews, 2010). Nonetheless, Dps-like proteins and ferritins have major differences that set them apart. For example, ferritins form 24-mers with an O symmetry while Dps monomers form 12-mers with a T symmetry. A fifth helix (helix E) that exists in ferritins is missing from the Dps proteins, suggesting that its absence may be responsible for the formation of the 12-mer rather than the 24-mer. It has been proposed that 24-mers evolved from a common dodecameric protein that became adapted to a number of functions giving rise also to the Dps family (Andrews, 2010; Grant et al., 1998; Peña & Bullerjahn, 1995).

The dodecamer forms a spherical cage with a hollow interior cavity in the middle. The empty space plays a key role in iron storage, being able to accommodate approximately up to ~500 iron atoms. The external diameter of the nano-cage is 80-90 Å, while the interior is 40-50 Å. The surface of the exterior, as well as the interior, are negatively charged (Grant et al., 1998; Haikarainen & Papageorgiou, 2009). An exception is HP-NAP, a member of the Dps family, which has a positively charged exterior. This is possibly related to the function of HP-NAP as an adhesin (Zanotti et al., 2002).

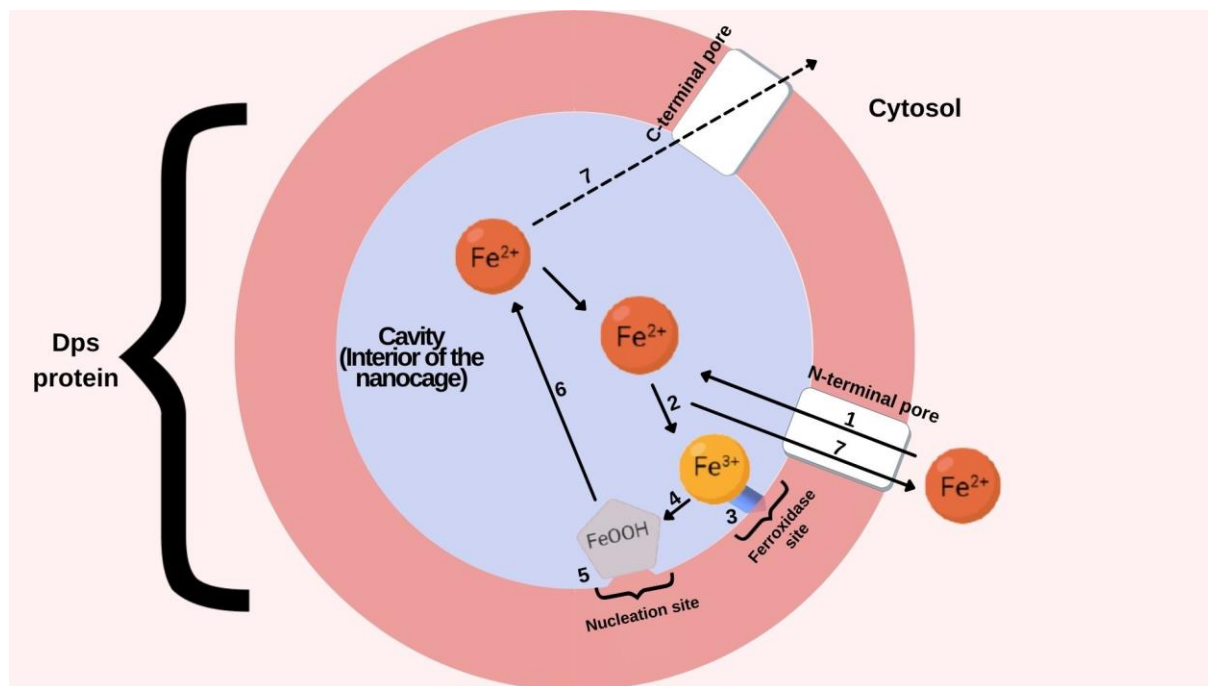


**Figure 1.** (A) Dodecameric structure with T symmetry of the *E. coli* Dps protein (B) Ribbon diagram showing the four- $\alpha$ -helix bundle of antiparallel helices of the monomer. Visualized and edited with Chimera (Pettersen et al., 2004)

## 1.4 Iron Incorporation

### 1.4.1 Iron Entry Mechanism

As already mentioned, Dps proteins are valuable and effective proteins able to protect against oxidative stress, by incorporating iron in a special compartment within their structure. A total of five steps are needed for iron incorporation, as depicted in Figure 2: a) iron entry in the dodecamer structure, b) iron binding to the ferroxidase center, c) oxidation of iron in the ferroxidase center, d) nucleation and e) mineralization.



**Figure 2.** Iron entry, oxidation, mineralization, and reutilization within the ferroxidase center in a Dps protein. Figure created with Canva, Polychronis Tatsis.

In contrast to ferritins which have 24 subunits that form an octahedrally (432) symmetrical shell, Dps dodecamers have a 23-symmetry. This type of symmetry creates four threefold axes that pass through the protein resulting in eight trimeric interfaces around the protein. These are formed at the end of the C-terminus or at the end of the N-terminus of each monomer, leading to the creation of channels that act as connecting pores between the interior and the exterior of the protein. These pores at the N-terminus are hydrophilic and negatively charged. This indicates a high resemblance with the iron-uptake pores of the ferritins, and accordingly, they have been proposed to be the route of iron insertion within the Dps and Dps-like proteins (Macedo et al., 2003; S. M. Williams et al., 2017).

In contrast to the N-terminal pores, C-terminal pores are highly hydrophilic and show a wider variability among Dps proteins than the N-terminal ones. Having a smaller size and due to their hydrophilic nature, the C-terminal pores are not considered to be the main pathway for cations. However, they might play a role as being part of the auxiliary route for cation passage after structural rearrangements (Kauko et al., 2004).

#### 1.4.2 Ferroxidase site

In numerous structures of the above-mentioned proteins, iron has been found to be bound in a site called the ferroxidase site. The dodecamer of the Dps-like proteins contains twelve

ferroxidase sites. All bacterial Dps proteins contain a ferroxidase site between each dimer, with the exception of *L. lactis* Dps (Stillman et al., 2005). Furthermore, in some Dps-like proteins, a conserved water is modeled instead of the iron in the ferroxidase site. According to Haikarainen & Papageorgiou et al (2009) “This water has an atypical distance to the iron at the ferroxidase center (about 1 Å longer than expected), which actually corresponds well with a typical 3.1 Å distance of l-oxo bridged irons. Due to the nature of this second iron as a catalytical intermediate, it is possible that water modeled in some Dps structures is actually a low occupancy iron atom.”

### 1.4.3 Iron core formation

As previously mentioned, iron is incorporated into the core of the Dps protein. The iron core of both ferritins and Dps proteins has been studied. Techniques such as X-ray absorption spectroscopy (XAS), Mossbauer spectroscopy, and polarized single-crystal absorption microspectrophotometry have been utilized for the characterization of the iron core. According to them, two different types of cores have been identified: i) a native core and ii) an *in vitro* core.

The native core is present in the “as purified” protein and no iron is inserted to the protein. This core contains tens of iron atoms and also other metals such as Zn, Cu, Cr, Mn, Co, Ni, and Mo. Unlike the native core, the *in vitro* loaded core of Dps-like proteins contains more iron atoms, up to 500 iron atoms. This core is formed after the addition of iron into the protein and its structure is less regular in comparison with the native core, which appears to be coordinated in an octahedral manner. The *in vitro* loaded core is characterized as a mixture of octahedrally and tetrahedrally coordinated ferrihydrite (Kauko et al., 2006; Yamamoto et al., 2002).

## 1.5 DNA Binding

As mentioned above, numerous, yet not all, Dps-like proteins use DNA binding as a mechanism to protect the bacteria from oxidative stress. This DNA-Dps protein complex forms a two-dimensional highly ordered honeycomb-like structure (Almiron et al., 1992; Haikarainen & Papageorgiou, 2009). This structure is ultra-stable and does not disintegrate even under high temperatures (up to 100°C) and even after treatment with detergents and/or solvents. Taking into consideration the fact that DNA and Dps are both negatively charged, it

has primarily been suggested that the binding occurs thanks to the positively charged residues of the N-terminus. However, several studies have suggested that these residues of the N-terminus are not directly responsible for the binding. It has been further explained that the interaction is the result of multiple ion bridges formed between DNA and Dps by divalent cations (Frenkiel-Krispin et al., 2001). Despite the importance of all of the above, the exact mechanism for the Dps-DNA binding is not entirely understood (Haikarainen & Papageorgiou, 2009).

After Dps-like proteins bind to the DNA, two models of Dps-DNA biocrystals are formed. The first model proposes that the DNA helix binds to the channels between three adjacent dodecamers in a lattice that is characterized as “pseudo-hexagonal”. The DNA-binding is mediated by three lysines at the disordered N-terminus of each of the monomers. As a result, each channel at the hexagonal lattice is lined by nine lysine residues, three from each monomer (Dadinova et al., 2019).

The second model proposes that the insertion of DNA to Dps is considered to be causing no disturbance in the intraplanar lattice spacing of the biocrystals. This revealed that DNA is packed within the biocrystals formed by the DNA and the Dps and also that both of them are stacked in alternating layers (Dadinova et al., 2019; Haikarainen & Papageorgiou, 2009).

## **1.6 Applications of the Iron Core**

The iron core has major applications in drug delivery and by extension in medicine. To be exact, iron core has contributed to nanomedicine (Herrmann et al., 2011), especially in the creation of new nanoparticles. Nanomaterials with a magnetic core have great potential as therapeutic and/or diagnostic tools in medical applications. Cancer treatment is the number one field in which magnetic nanoparticles can be utilized (Gupta & Gupta, 2005; Singh et al., 2010).

Regarding the diagnostic aspect, magnetic nanomaterials can be detected by two techniques: i) magnetic resonance imaging (MRI) and ii) X-ray computed tomography (CT). In the field of therapeutics, nanoparticles with an iron core can be utilized as drug carriers for cancer treatment, controlled by external stimuli (magnetic field), or as thermotherapeutic probes under radial high-frequency fields (Chou et al., 2010; Ito et al., 2006; Shieh et al., 2005; P. C. Wu et al., 2007).

One of the most widely used nanoparticles are golden nanoparticles with an iron core (Fe@Au) (Zhou et al., 2001), used as cancer-selective drug carriers, without damaging healthy cells and organs. Fe@Au nanoparticles were designed by Zhou et al. and take advantage of the magnetic susceptibility of zero valent iron cores to guide the nanoparticles. Au is used as a protective medium for the iron core, preventing its oxidation from the environment (Cho et al., 2005; Zhou et al., 2001).

## 1.7 Bacterial Infections & Antibiotic Resistance

From the very first discovery of penicillin back in 1928 when Alexander Fleming made a breakthrough, all the way to the 20th century, the rapid progress of life sciences gave the impression that bacterial infections will be easily dominated (Vouga & Greub, 2016). In the 21st century, antibiotic resistance started to become a major health issue as a result of antibiotic overuse and the ability of some of the microbes to resist these drugs. Antibiotic resistance is such an emerging global threat, that according to Avershina et al., 2021 in 2050 this issue will cause more deaths than all the types of cancer combined together. A study published in 2022 highlights the significance of antimicrobial resistance, stating that in 2019, antibiotic resistance lead more people to death than HIV/AIDS or malaria (Murray et al., 2022).

Several studies include the bacteria of the genus *Streptococcus* within their lists of emerging bacterial pathogens (*Streptococcus bovis* group) (Vouga & Greub, 2016) or of bacteria in need of new antibiotics (*Streptococcus pneumoniae*, as medium priority pathogen for new antibiotics) (retrieved from <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>, assessed February 2017).

More specifically, resistant determinants of *S. suis* for compounds such as tetracyclines, macrolides, aminoglycosides, chloramphenicol, antifolates, streptothricin, and cadmium salts have been identified. *S. suis* acts as a resistance reservoir that may attribute to the spread of antibiotic resistance genetic elements to other streptococcal human pathogens such as *S. pyogenes*, *S. pneumoniae*, and *S. agalactiae* (Palmieri et al., 2011). Another recent study conducted by Lunha et al., 2022 demonstrated that there is a slight decrease in *S. suis* strains susceptible to the currently used antibiotics raising awareness for an upcoming resistance problem in cases where this bacterial species is implicated.

In conclusion, it is clear from the above-mentioned studies that the formation of the iron core plays an important role in the protection of bacterial cells from oxidative stress. Pores are proposed to work as the entrance of iron, so blocking these gates is of high importance for interfering with the cell viability. Antibiotic resistance is an emerging problem, so pore blocking could be used as an alternative therapeutic approach for the treatment of the infections.

## 2 Aims

The aims of this study are as follows:

- To discover crystallization conditions for the 6xHis-tagged Dpr to avoid the 6xHis-tag removal step by protease cleavage.
- To identify potential N-terminal blockers by computational protein ligandability investigation.
- To provide insights into the protein-ligand interactions in the N-terminal pores.



## 3 Materials & Methods

This project was divided into two parts: Protein crystallization & Structure Determination (3.1) and Computational/Ligandability Part (3.2).

### 3.1 Protein Crystallization & Structure Determination

This part was carried out as summarized in Fig. 3 in the Results section.

#### 3.1.1 Bacterial Culture & Protein Purification

The recombinant bacterial strain of *E. coli* with pSSUD2 plasmids expressing the 6xHis-tagged *S. suis* Dpr (*ssDpr*) was utilized. 700 mL of *E. coli* bacterial culture was centrifuged at 300 x g for 15 min at 4 °C. Subsequently, the bacterial pellet was stored at -84 °C overnight. The following day the bacteria were left to thaw in an ice bath for 30 min. The bacteria were suspended in 10 mL 50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.5 NaCl, 20 mM imidazole. Following that, lysozyme was added to a final concentration of 0.4 mg/mL along with 20 µg/mL DNase, 1 mM MgCl<sub>2</sub>, 2 mM PMSF (phenylmethylsulfonyl fluoride). The solution was incubated for 30 min at room temperature (RT) and then sonicated with Qsonica Sonicator Q700 (Thermo Fisher Scientific, Waltham, MA USA) at 6 x 15 s pulses at 30 s intervals with medium speed and centrifuged for 30 min, at 16,000 x g at 4 °C. Filtration for debris and lysate separation was conducted with a 2 µm filter attached to a Luer Lock syringe.

#### 3.1.2 Protein Purification & Concentration Determination

Truncated 6xHis-tagged Dpr protein with the first 7 N-terminal amino acids missing, was initially provided by our collaborator (Sauli Haataja, Institute of Biomedicine, University of Turku). pSSUD2 plasmid within *E. coli* strains was utilized for protein production. Two protein purification schemes were carried out that led to two protein batches (batch one and batch two). Protein purification was different for batch one and batch two.

#### 3.1.3 Batch One

Protein purification performed via ÄKTAprime plus, utilizing a PROTEINDEX™ HiBond™ Ni-NTA Agarose 6FF, Prepacked Cartridge (Marvelgent Biosciences, Canton, MA, USA). Equilibration and washing were performed with buffer 50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.5 M NaCl, 20 mM imidazole. The same buffer was used for elution, but with a linear gradient of 20 to

500 mM imidazole, receiving 20 ml of 1 ml fractions. Following purification, buffer exchange was conducted. Phosphate buffers are not preferred for crystallization experiments, due to their ability to form salt crystals. The initial buffer of 50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.5 NaCl, 20 mM imidazole, was gradually changed to HEPES ((4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)) 10 mM, pH 7.0, NaCl 150 mM, NaN<sub>3</sub> 0.002 % w/v. After 5 cycles of using Millipore Amicon ultra centrifugal filter units (Amicon, Miami, Florida, USA). Flow-through was kept in case of a ripped filter. 2 mL of Dpr were obtained. NanoDrop™ Microvolume Spectrophotometer (Thermo Fisher Scientific, Waltham, MA USA) was used for protein quantification, setting the buffer as blank. After retrieving the results, Dpr concentration was determined to be ~1.2 mg/mL. Due to the low amount of protein and its low concentration, it was decided that 6xHis-tag would not be removed with thrombin as initially planned. Instead, crystallization with the tag would be performed. Protein aliquots were stored at 4 °C.

#### 3.1.4 Batch Two

Protein purification was conducted with HisPur™ Ni-NTA Spin Columns, 3 mL (Thermo Fisher Scientific, Waltham, MA USA). The column was equilibrated at RT and sample preparation was conducted by admixing the crude protein extract with an equal volume of the equilibration buffer. Then, the bottom tab of the column was removed by applying slight pressure and gently twisting the tab and the column was placed into a 50 mL centrifuge tube. Centrifugation at 700 x g for 2 min was performed for the storage buffer removal. After this step, the flow-through was discarded and 6 mL of equilibration buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.5 M NaCl, and 20 mM imidazole) was used to equilibrate the column. Centrifugation at 700 x g for 2 min was performed for buffer removal. The column bottom plug was attached again to the column and the crude protein extract was added to the column and left at an orbital shaker for 30 min at RT. Next, the bottom plug was removed again and the column was put for centrifugation at 700 x g for 2 min, and the flow-through was collected in a centrifuge tube. Furthermore, the resin within the column was washed with 6 mL of wash buffer 50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.5 M NaCl, and 20 mM imidazole and centrifuged at 700 x g for 2 min. The fraction was collected in a centrifuge tube and this step was repeated two additional times, collecting each time the fraction in a separate centrifuge tube labeling each tube. Protein elution was monitored by Coomassie Plus (Bradford) Assay Reagent (Cat. No. 23238) and the purity was also visualized by SDS-PAGE.

### 3.1.5 Storage

Purified 6xHis-tagged protein was stored at 4 °C in labeled aliquots in both cases after purification for further use.

### 3.1.6 Crystallization Screenings

For the first batch of protein, three crystallization screens were employed: SG1™ Screen HT-96 (Molecular Dimensions, Sheffield, England, UK), JCSG-plus HT-96 (Molecular Dimensions, Sheffield, England, UK), and Index HT (Hampton Research, Aliso Viejo, California, USA).

Swissci MRC Crystallization 96-Well 2-Drop Plates (Molecular Dimensions, Sheffield, England, UK) were utilized for the screening procedure. Crystals were grown by the hanging-drop vapor-diffusion method in all of the screens and in all of the scale-up procedures. Each plate represents an individual screening process. Screening reagents that were previously stored at 4 °C, were left to RT for 10 min prior to the initiation of the screen. Each MRC plate was labeled appropriately on the side according to the screening name and protein name. Wells were filled with 40 µL of screening conditions using a multichannel pipette. After all volume depositions, cling film was used to cover the plates in order to avoid potential evaporation.

Following this, appropriate layouts for each MRC plate were created by the RockMaker software (Formulatrix, Bedford, Massachusetts, US,) and printed as a barcode sticker which was attached to the side of each MRC plate. Next, each plate was placed on mosquito® LV (SPT Labtech, Melbourn, UK) on the right panel. Micro-reservoir plastic stripes were used as protein deposition wells on the left panel of mosquito LV. Each well of the micro strip was filled with 1.75 µL of ssDpr and magnetic metallic holders were placed on top of the micro strip in order to secure its position. After the completion of the run, each plate was sealed with an adhesive plastic cover sheet making sure that the top surface of the plate is smooth after the application of the cover and bubble-free. Following this, the plates were placed within the Formulatrix ROCK IMAGER RI 54 incubator at 20 °C (Formulatrix, Bedford, Massachusetts, USA).

### 3.1.7 Scale-up & crystallization optimization

An *ssDpr* protein aliquot previously stored at 4 °C was left in an ice bath. Pre-labeled 24-well Linbro plate (Jena Bioscience, Thüringen, Germany) was utilized and Dow Corning® high-vacuum silicone grease (Sigma-Aldrich, St. Louis, Missouri, USA) was applied to the circumference of every well rim using a 10 mL syringe. Plasticine was applied around the corners of the plate base. Ranging concentrations of buffer, precipitate (polyethylene glycerol (PEG)), salt, and H<sub>2</sub>O were deposited within each well. Following this, the plate was placed on a plate shaker for 5 min at quick shaking to let the chemicals to mix. After the shaking was completed, one drop of 2 µL *ssDpr* solution was deposited on the surface of a clean coverslip with where the dust had been previously removed from its surface with a brush. Next, 2 µL of the reservoir solution was deposited and mixed with the protein drop (ratio 1:1), taking care at the same time to avoid any bubble formation. The cover slip was then inverted and placed on the rims of the well, applying gentle pressure in order to seal the well with the grease and to create a closed system. After adjusting all the coverslips in a tangent manner in order to avoid overlapping, the plate was covered by its lid and placed in a Porkka incubator (Porkka, Lahti, Finland) at 16 °C.

### 3.1.8 IceBear

Wells of the screening plates within the Formulatrix incubator were observed via the IceBear software (Biocenter Oulu and University of Oulu, Faculty of Biochemistry and Molecular Medicine, Oulu, Finland) (Daniel et al., 2021). Images were taken via a micro camera in specific time intervals based on the Fibonacci sequence and uploaded to a cloud database accessible by the IceBear software.

### 3.1.9 Stereoscope Visualization & Crystal Collection

Wells were observed under an SZ-PT stereoscope (Olympus, Tokyo, Japan) and those with crystals of appropriate size, were marked down and utilized for crystal collection and storage. Observation time was kept as short as possible to avoid large fluctuations of the temperature. Appropriate size CryoLoops on a CrystalCap Magnetic (Sheffield, England, UK, Molecular Dimensions) (0.1 – 0.2 mm diameter) attached on a magnetic wand, were utilized for crystal fishing. Cryoprotectant solutions with different glycerol concentrations depending on the conditions of the well were created and the fished crystals were dipped within the cryoprotectant drop prior to flash cooling them in bulk liquid nitrogen at 66-77 K. The

CryoLoops were placed in pucks of SPINE (SC3) Baskets (MiTeGen, Ithaca, NY, USA) which were subsequently packed in appropriately labeled cryoshippers (MVE SC4/2V; (Mve Biological Solution, Brentwood, TN, USA) and transferred to the X-ray synchrotron facility for data collection, where the temperature of the cryosystem for data collection was 100 K.

### 3.1.10 X-ray Diffraction

In-house MicroMax-007 HF (Tokyo, Japan, Rigaku) X-ray diffraction system was utilized for preliminary crystal characterization. The synchrotron facility at the German Electron Synchrotron (DESY; Hamburg, Germany) was utilized for full crystallographic data collection.

## 3.2 Computational/Ligandability Part

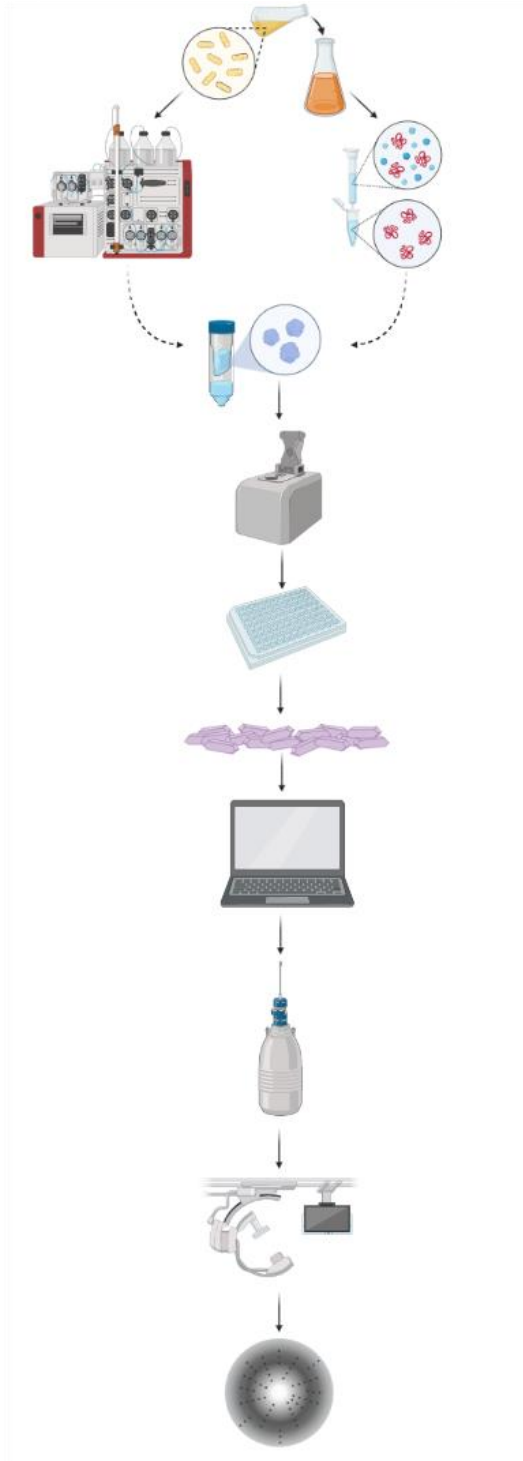
The *ssDpr* protein structure deposited in PDB (ID: 1UMN) (<http://doi.org/10.2210/pdb1UMN/pdb>) was used for this part. This structure was produced using crystals of protein that had the 6xHis tag removed, thus crystallized without the presence of the 6xHis tag.

LiGANN open-source software (Goodfellow et al., 2014; Jiménez et al., 2018; Skalic et al., 2019; Zhu et al., 2017) was utilized with default settings for structured-based *de novo* ligand library creation. Cavity-detection guided Blind Docking (CB-Dock) open-source software (Cheng et al., 2017; Liu et al., 2020) was utilized for individual ligand docking. Chimera (Pettersen et al., 2004) software was utilized for depiction and analysis purposes both for the protein-ligand complex and for the ligand standalone.

## 4 Results

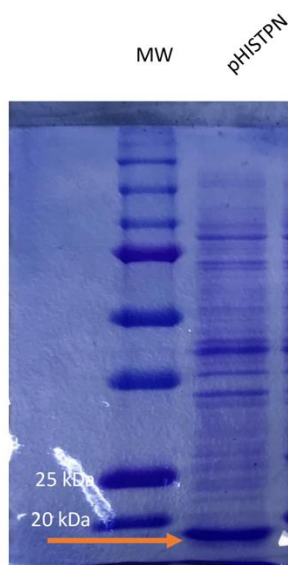
### 4.1 Protein Purification, Concentration Determination & Storage

Figure 3 below depicts the pipeline followed for this part.

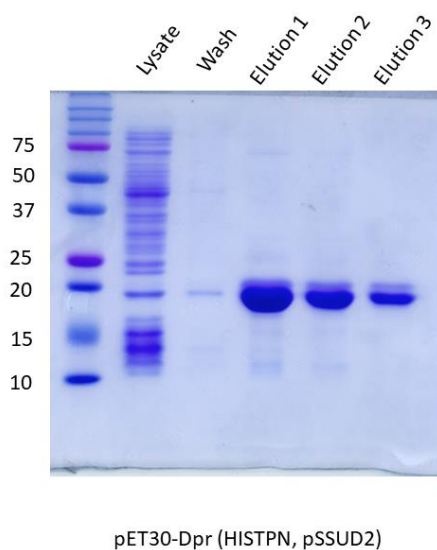


**Figure 3.** Protein Crystallization & Structure Determination Pipeline.

Figure 4 depicts the purity of the protein lysate, as indicated by the bands at 20 kDa. Figure 5 depicts the purity of the protein and as seen by the bands in the last three (3) columns. There is no significant amount of by-products or other proteins. Band intensities are proportional to the protein concentration. Moreover, Table 1 presents the concentrations determined for each batch.



**Figure 4.** SDS-PAGE of lysate purity of the first batch.

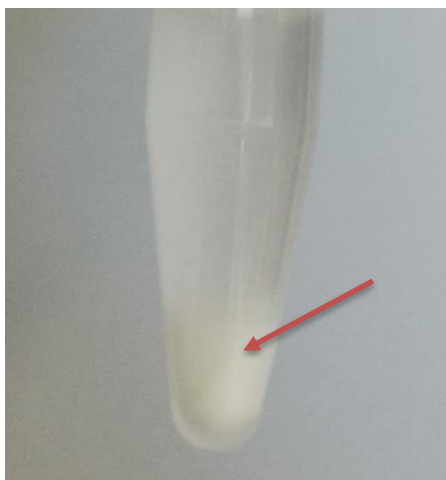


**Figure 5.** Protein purity of the second batch after purification as a result of SDS-PAGE. Molecular weights of the standards are shown on the left.

**Table 1.** Protein concentration for each batch

<b>Batch</b>	<b>Concentration</b>	<b>Total volume</b>
First	~ 1.2 mg/mL	~ 2 mL
Second	~ 10 mg/mL	~ 5 mL

During storage, Dpr started to precipitate in the aliquots a day after the screenings, as depicted in Figure 6.

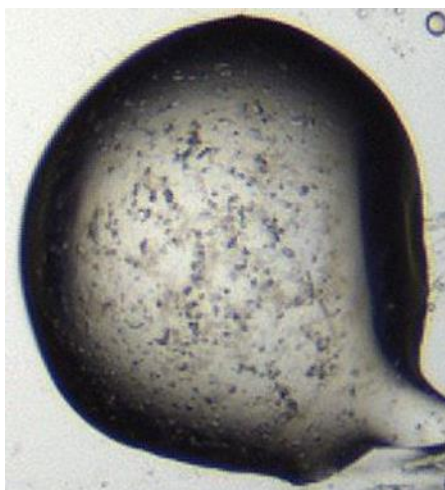
**Figure 6.** Precipitation in the second protein batch.

## 4.2 Protein Crystallization & Structure Determination

### 4.2.1 First Batch

Crystals were grown in condition B05.1 of the MD1-88 SG1 screen (Figure 7), corresponding to condition 1 as depicted below. All conditions that yielded crystals are listed in Table 2.



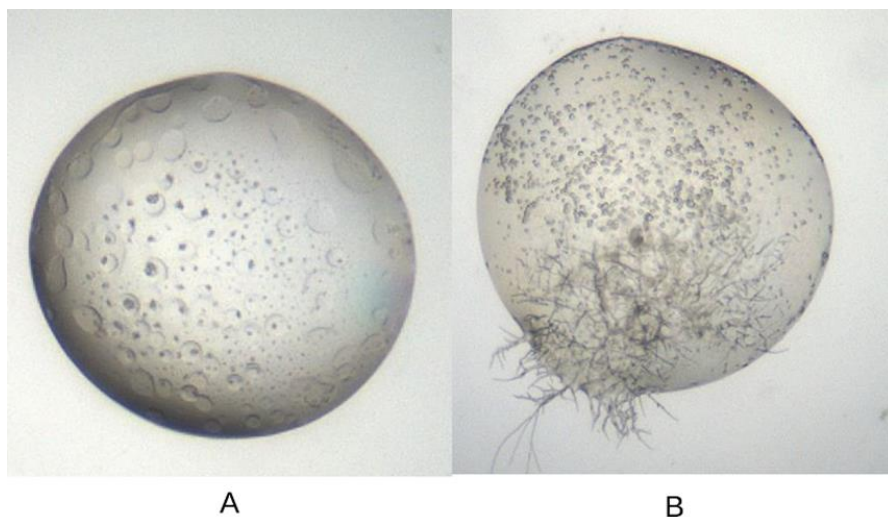


**Figure 7.** Crystal growth in screen MD1-88 SG1 from Molecular Dimensions, condition 1.

**Table 2.** Crystallization conditions where crystals developed

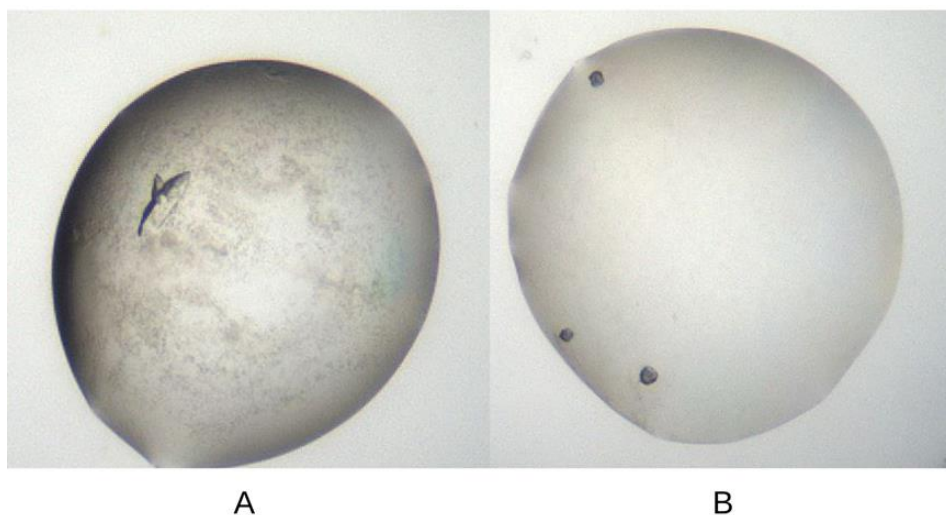
Crystallization screen	Well position of the screening condition	Conditions	Condition No
<b>First Batch</b>			
MD1-88 SG1	B05	0.2M Calcium chloride dihydrate, 20% w/v PEG 3350	1
MD1-88 SG1	A10	0.2M Lithium sulfate, 0.1M Tris 8.5, 30% w/v PEG 4000	2
MD1-88 SG1	F10	0.2M Magnesium acetate tetrahydrate, 20% w/v PEG 3350	3
MD1-40 JCSG plus HT-96	E11	0.16M Calcium acetate hydrate, 0.08M Sodium cacodylate, pH 6.5, 14.4% w/v PEG 8000, 20% v/v glycerol	4
MD1-40 JCSG plus HT-96	D10	0.2M Calcium acetate hydrate, 0.1M Sodium cacodylate, pH 6.5, 40% v/v PEG 300	5
<b>Second Batch</b>			
MD1-88 SG1	E07	0.2M Magnesium chloride hexahydrate, 20% w/v PEG 3350	6
MD1-40 JCSG plus HT-96	B08	0.2M magnesium chloride hexahydrate, 0.1M Tris, pH 7, 10% w/v PEG 8000	7
MD1-40 JCSG plus HT-96	G04	0.2M TMAO, 0.1M Tris, pH 8.5, 20% w/v PEG 2000 MME	8
MD1-88 SG1	F08	0.2M Potassium nitrate; 20% w/v PEG 3350	9
MD1-88 SG1	F02	0.2M Sodium acetate trihydrate; 0.1M Sodium HEPES 7.5; 25% w/v PEG 3350	10

At the same screen, there were also other wells (A10.1 and F10.1) presenting crystal-like images such as those depicted in Figure 8 below.



**Figure 8.** Crystal growth in wells A10.1 (A) and F10.1 (B) in crystallization screen MD1-88 SG1, corresponding to conditions 2 and 3, respectively, in Table 2.

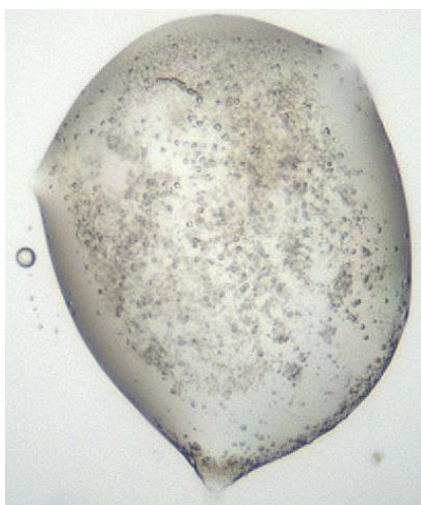
Screen MD1-40 JCSG plus HT-96 provided two hits (E11.1 and D10.1), depicted in Figure 9 below.



**Figure 9.** Crystal growth in wells E11.1 (A) and D10.1 (B) in crystallization screen MD1-40 JCSG plus HT-96, corresponding to conditions 4 and 5, respectively, in Table 2.

#### 4.2.2 Second Batch

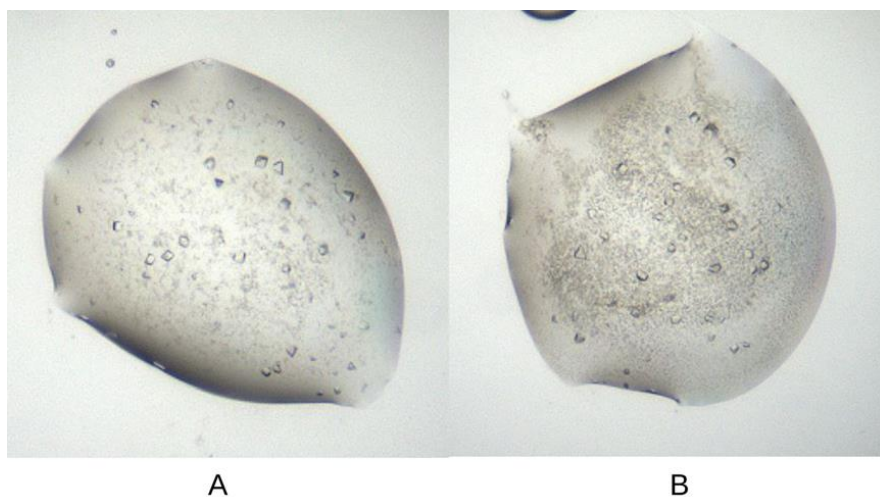
Crystallization screens were set up with two concentrations of Dpr. Wells to the first position correspond to half of the initial concentration, i.e., ~5 mg/mL and wells to the second position correspond to the initial concentration, i.e., ~10 mg/mL. Screen MD1-88 SG1 gave only one hit (Figure 10), in well E07.2 with the conditions of 0.2M magnesium chloride hexahydrate, 20% w/v PEG 3350. The hit was examined visually and it was decided that it resembles microcrystalline precipitation, so no further examination of it was conducted.



**Figure 10.** Microcrystalline formation in well E07.2 in crystallization screen MD1-88 SG1, Molecular Dimensions, corresponding to condition 6.

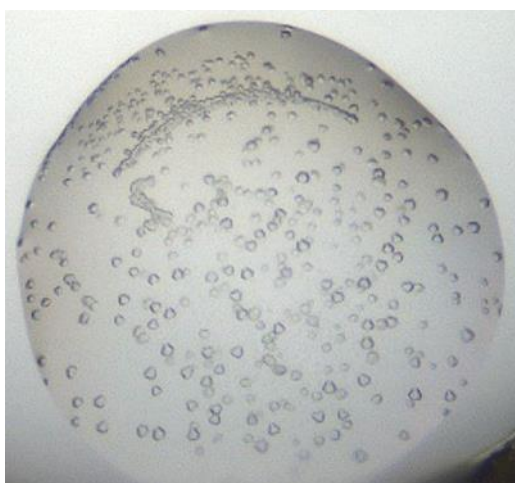
Screen MD1-40 JCSG plus HT-96 provided hits in wells B08.1 and B08.1, which appeared to have visible crystals after Day 8. Crystal shapes are depicted below in Figure 11.

The plate with varied pH and PEG did not provide appropriate crystals.



**Figure 11.** Crystal growth in wells B08.1 and B08.2 in crystallization screen MD1-40 JCSG plus HT-96, Molecular Dimensions, corresponding in condition 7.

The well G04.1 appeared to have crystal formations as shown below in Figure 12.



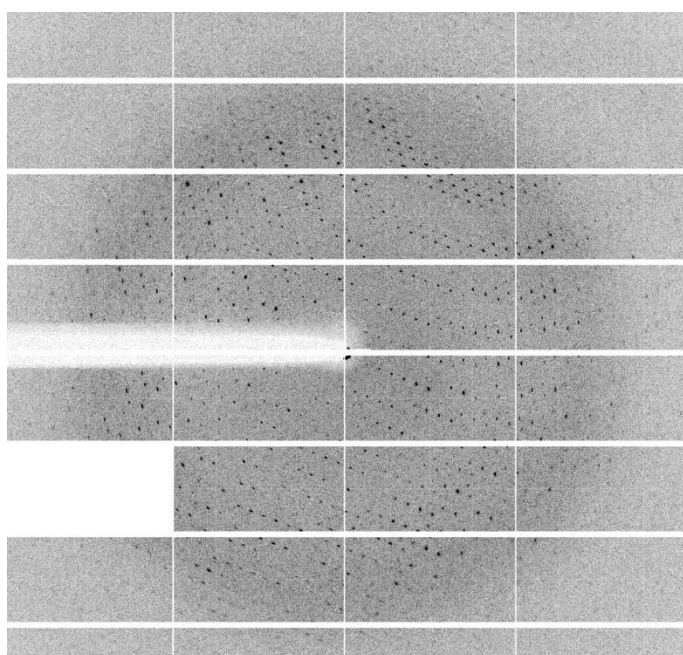
**Figure 12.** Crystal growth in well G04.2 in crystallization screen MD1-40 JCSG plus HT-96, Molecular Dimensions, corresponding to condition 8.

Crystals in conditions 9 and 10 were visible after Day 21, meaning that they could not get visualized with IceBear scheduled inspections. Instead, crystals were found to have grown via visual inspection under the stereoscope.

### 4.2.3 X-ray data collection

A total of 25 crystals from conditions 7, 9, and 10 (Table 2) were mounted in cryoloops and shipped to EMBL-Hamburg for testing and possible data collection. Data were collected from crystals of conditions 9 and 10. Condition 7 did not give any diffraction data.

X-ray diffraction data to 2.2 Å were obtained on the P13 beamline at the PETRA III synchrotron facility (Figure 13). This diffraction pattern resulted from a crystal grown in condition 9 (Table 2) and the structure is yet to be determined. The crystal used is depicted in Figure 14. The crystal from condition 10 resulted in diffraction data up to 3.63 Å resolution.



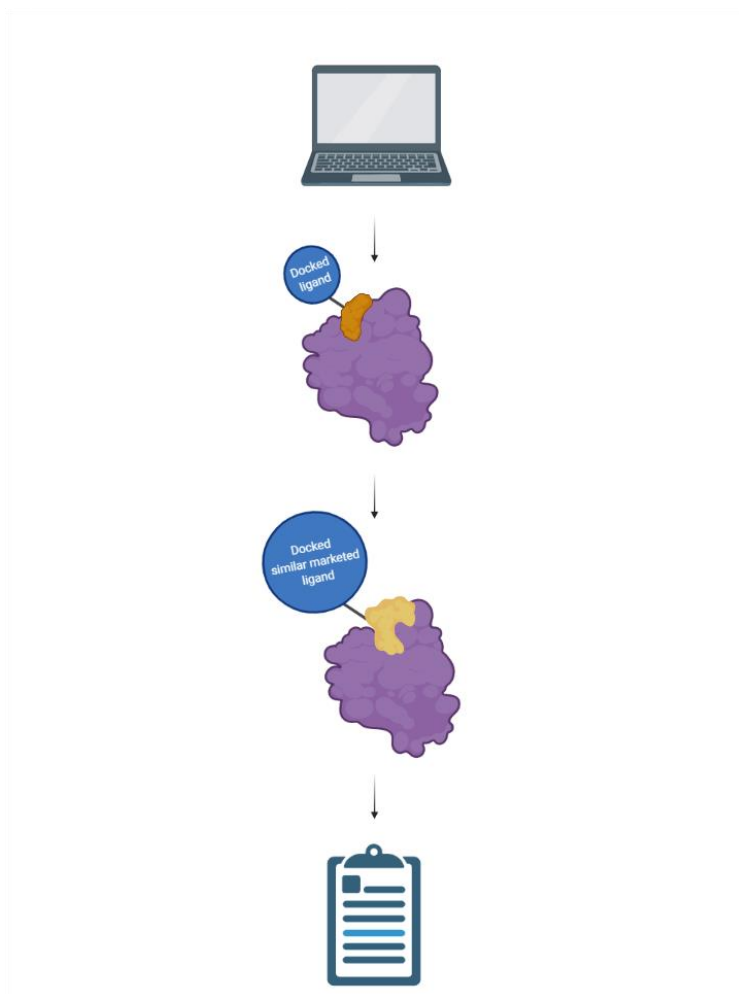
**Figure 13.** Representative X-ray diffraction pattern of a Dpr crystal from condition 9. The resolution is 2.2 Å.



**Figure 14.** Dpr crystal (~40  $\mu\text{m}$  size) corresponding to condition 9 that gave the best diffraction pattern and the 2.2 Å full dataset. The loop size is 100  $\mu\text{m}$ .

#### 4.3 Computational/Ligandability Part

Apart from the laboratory part results, information was also obtained for the *in silico* part of this project. Figure 15 below presents the steps followed for this scientific part.



**Figure 15.** Computational/Ligandability Part Pipeline.

#### 4.3.1 Ligand Library Creation

Initial creation of novel ligands for the binding pocket resulted in the creation of 82 potential ligands (Table 3). After individual docking at the binding site, all ligands with a docking score better than -7 were pooled, resulting in a new library of 20 ligands, as depicted in Table 4. Initial screening of those 20 ligands in the ZINC database, gave 2 hits, concluding that only two of these ligands are actually purchasable. Both of these ligands contain two aromatic rings (Figure 16). The other 18 are not purchasable and cannot be used in order to validate in the lab the *in silico* binding, e.g., via protein crystallization or biological assays. Another search in ALDRICH MARKET SELECT (<https://www.aldrichmarketselect.com/>, accessed April 2022) was run, in order to double check if these 20 ligands can be purchased. None of the 20 ligands was found in the database.

**Table 3.** De novo ligand library

No	Smiles	Docking Score (kcal/mol)
1	<chem>CCCCC(CC)CNC(C1CC1)C1CN(C(=O)OC(C)(C)C)C1</chem>	-6.1
2	<chem>CCN(CC)CC(C)NCCc1csc(C)n1</chem>	-5.3
3	<chem>CCCNC(=O)CCC(=O)N(CC1CCCC1)C(CC1CCCC1)c1cccc1</chem>	-4.3
4	<chem>COCC(O)CN1CCC2=C(c3cccc3)CCC(=O)C2C1</chem>	-7.2
5	<chem>FC1(F)C=CCC1</chem>	Docking in incorrect cavity
6	<chem>NC(=O)C(F)(F)C(F)(F)F</chem>	-5.1
7	<chem>COCC(COC)NCC(c1cccc1)c1ncc[nH]1</chem>	-6.3
8	<chem>CCCC(=O)N(C)Cc1ccco1</chem>	-5.9
9	<chem>Fc1ccsc1CN1CCC1</chem>	-5.1
10	<chem>CC(C)NC(=O)CN1CCN(CCCN(Cc2cccn2)Cc2cccn2)CC1Cc1cnn(C)c1</chem>	-6.9
11	<chem>CC#CC(=O)N(C)Cc1ccco1</chem>	-6.1
12	<chem>C#CCOCc1cccc(NCc2coc3ccccc23)c1</chem>	-7.5
13	<chem>O=C(c1ccco1)N1CCC1</chem>	-5.6
14	<chem>O=C(Br)C(F)(F)c1ccco1</chem>	-5.6
15	<chem>CCOCC(=O)NCCn1c2ccccc2c2ccccc21</chem>	-7.0
16	<chem>CCN(CC)CCNC(=O)C1CCCN(CCS2nnc(Cc3cccn3)n2Cc2ccco2)C1</chem>	-6.9
17	<chem>COCC(O)CN1CCc2[nH]c3ccccc3c2C1</chem>	-6.9
18	<chem>O=C(Nc1ccc(Cl)cc1)C(Sc1nnc(C2CC2)n1Cc1cccc1Cl)c1cc2ccccc2[nH]1</chem>	-8.6
19	<chem>C#CCNCc1cn(-c2ccccc2)nc1-c1cccs1</chem>	-6.8
20	<chem>C#CCNCc1cccc(OCc2cccs2)c1</chem>	-6.5
21	<chem>N#CC(=O)N1CCc2cccn1c21</chem>	-6.5
22	<chem>O=C(C1C=CCC1)N1CCC1</chem>	-5.5
23	<chem>O=C(O)C1(C(F)(F)F)CC1(F)F</chem>	-5.5
24	<chem>COCCNC(=N)NCc1cccc1-n1ccn1</chem>	-6.6
25	<chem>C#CCCCC(=O)N(CC1CCC2(C1)CC(C(C)C)C2)C(C)c1cccn1</chem>	-6.5

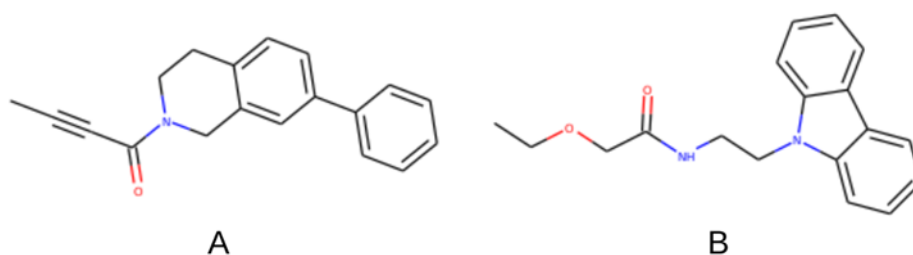


26	<chem>CCCNC(=O)NCCN1CCN(CCCn2c(-c3ccccc3)nc3ccccc32)CC1C(C)C</chem>	-7.9
27	<chem>Cc1ccc(C(=O)NCC(c2ccc(Cl)cc2)c2nc3ccccc3n2CCCC(C)(C)C)cc1C</chem>	-8.4
28	<chem>CC(C)N1CCN(CCCNC(=O)N(CCCc2ccccc2)Cc2ccccc2)CC1c1nccn1C</chem>	-7.5
29	<chem>O=C(O)C(F)c1ccco1</chem>	-5.4
30	<chem>Fc1ccsc1C(F)(F)C(F)F</chem>	-5.8
31	<chem>O=C(N1CCCC1)C(F)(F)F</chem>	-5.5
32	<chem>O=C1NCC2CC12</chem>	-3.8
33	<chem>CC(CNC(=O)NC1CCN(C2CC2)C1)CN1C(=O)c2ccccc2N2CCC(Cc3ccccc3)N12</chem>	-7.8
34	<chem>O=C(N=S)N1CCC1</chem>	-3.8
35	<chem>COCCNC(=N)NCc1nnc(-c2ccccc2)s1</chem>	-5.9
36	<chem>N#CCCOc1ccccc1-c1ccc(-c2ccccc2)cc1</chem>	-7.5
37	<chem>CCC=CC(=O)N1c2ccccc2CC1C</chem>	-6.7
38	<chem>FC1CCC1(F)c1ccco1</chem>	-5.1
39	<chem>CCN(CCN1CCN(C(=O)CCCc2nnc(-c3ccccc3)o2)CC1)Cc1cccn1</chem>	-7.2
40	<chem>CN1CC(Nc2ncccc2C#N)CC1(C)C</chem>	-6.5
41	<chem>CNCCSCC(=O)N1N=C(c2ccco2)CC1C</chem>	-6.3
42	<chem>C=CCCCCN(C)C(=O)CC12CC3CC(C1)CC(C(F)(F)F)(C3)C2c1ccnc1</chem>	-7.2
43	<chem>CCNC(=O)CCN1CCCC1Cn1ccn1</chem>	-5.9
44	<chem>Cn1[nH]ccc1=NC(=O)C1CC1(F)F</chem>	-6.3
45	<chem>Cc1ccc(NC(=O)C(C)SCc2nc(CC(C)C)no2)c(C(C)C)c1NCCc1cccc(Cl)c1</chem>	-7.5
46	<chem>COCC(C)(O)CNCc1ccccc1-n1ccn1</chem>	-6.2
47	<chem>CNCC1CCCC1NCc1cccn1</chem>	-5.9
48	<chem>O=C(O)c1ccnc1Br</chem>	-4.9
49	<chem>CC(CCN(C)C)N(C)Cc1ncccc1F</chem>	-5.5
50	<chem>CC#CC(=O)N1CCc2ccc(-c3ccccc3)cc2C1</chem>	-7.1
51	<chem>CNCC(C)CNC(=O)C(C)n1ccn1</chem>	-5.6
52	<chem>FC(F)(F)c1ccoc1</chem>	-4.2
53	<chem>O=C(CC(F)(F)F)N1CCC1</chem>	-4.8
54	<chem>CNCCC(=NNn1cccn1)C(F)(F)C(F)(F)F</chem>	-6.3

55	<chem>Fc1cccnc1l</chem>	-4
56	<chem>CNCCNC(=O)C1CCN(c2ccccc2)CC1</chem>	-6.5
57	<chem>Fc1ccsc1C(F)F</chem>	-4.4
58	<chem>CCSCCNc1cnns1</chem>	-4.3
59	<chem>CCN(Cc1ccsc1)C(C)CO</chem>	-5
60	<chem>C#CC(C)NCc1ccco1</chem>	-5
61	<chem>O=C(O)c1ccc[nH]1</chem>	-4.4
62	<chem>CCC=CCOc1ccccc1-c1cccc2ccccc12</chem>	-7.3
63	<chem>C#CCOCc1cccc(Cn2ccc3sccc3c2=O)c1</chem>	-6.6
64	<chem>CCCC(CC1CCCC1)C(=O)N(C)c1ccc(CNC(=O)C2CC2C(C)(C)C)O1</chem>	-7.1
65	<chem>FC1CCC1N=Cc1ccco1</chem>	-4.9
66	<chem>CC1CN(C(=O)OC(C)(C)C)C(C)CC1CCNCc1c[nH]nc1C(C)(C)C</chem>	-7.2
67	<chem>O=C(C=C(F)F)N1CCC1</chem>	-4.7
68	<chem>CCCC(C)NC(C)c1ccco1</chem>	-5.1
69	<chem>CC(F)(F)C(=O)c1ccco1</chem>	-4.9
70	<chem>CC1CCC1=NO</chem>	-4
71	<chem>CNCC#CCNC(C)C1OCCc2sccc21</chem>	-6.1
72	<chem>C=CCNC(=O)C(C#N)=c1sc(=CC=C2CCC(C)(C)C2)c(=O)n1CCCC(C)(C)C</chem>	-7.4
73	<chem>CCN1CCN(C2CCN(C(=O)CCCCn3c(S)nnc3-c3ccccc3)CC2)CC1c1ncc[nH]1</chem>	-7.5
74	<chem>O=C(C1CC1)N1CCC1</chem>	-4.2
75	<chem>FC1CCC1N1CCc2ccnc21</chem>	-6.1
76	<chem>COCCOC(=N)NC(c1ccccc1)c1nccs1</chem>	-6.
77	<chem>COCCNC(=S)NCc1ccccc1-n1cccn1</chem>	-5.9
78	<chem>Cc1ccc(CNC(Cc2ccc(C3CCCC3)cc2)c2nc(NC(=O)OC(C)(C)C)no2)cc1C#N</chem>	-8.2
79	<chem>CN(C)C(=O)c1ccco1</chem>	-4.4
80	<chem>O=C(O)C1CCcn2ccnc21</chem>	-5.5
81	<chem>CCN1CCN(CCCN2CCCC(c3nc(Cc4ccccc4)nn3-c3ccccc3)C2)CC1</chem>	-7.7
82	<chem>CCC=C(C)C(=O)NCc1ccco1</chem>	-5.6

**Table 4.** Ligands with descending docking scores from best to worst, keeping a score threshold of -7 kcal/mol.

No	Ligand No	SMILES	Score (kcal/mol)
1	18 (H bond)	<chem>O=C(Nc1ccc(Cl)cc1)C(Sc1nnc(C2CC2)n1Cc1cccc1Cl)c1cc2ccccc2[nH]1</chem>	-8.6
2	27	<chem>Cc1ccc(C(=O)NCC(c2ccc(Cl)cc2)c2nc3ccccc3n2CCCC(C)(C)C)cc1C</chem>	-8.4
3	78	<chem>Cc1ccc(CNC(Cc2ccc(C3CCCC3)cc2)c2nc(NC(=O)OC(C)(C)C)no2)cc1C#N</chem>	-8.2
4	26 (H bond)	<chem>CCCNC(=O)NCCN1CCN(CCCn2c(-c3ccccc3)nc3ccccc32)CC1C(C)C</chem>	-7.9
5	33	<chem>CC(CNC(=O)NC1CCN(C2CC2)C1)CN1C(=O)c2ccccc2N2CCC(Cc3cccn3)N12</chem>	-7.8
6	81	<chem>CCN1CCN(CCCN2CCCC(c3nc(Cc4cccn4)nn3-c3ccccc3)C2)CC1</chem>	-7.7
7	12 (H bond)	<chem>C#CCOCc1cccc(NCc2cc3ccccc23)c1</chem>	-7.5
8	28	<chem>CC(C)N1CCN(CCCNC(=O)N(CCc2cccn2)Cc2cccn2)CC1c1nccn1C</chem>	-7.5
9	36	<chem>N#CCCOc1cccc1-c1ccc(-c2ccccc2)cc1</chem>	-7.5
10	45	<chem>Cc1ccc(NC(=O)C(C)SCc2nc(CC(C)C)no2)c(C(C)C)c1NCCc1cccc(Cl)c1</chem>	-7.5
11	73	<chem>CCN1CCN(C2CCN(C(=O)CCCCn3c(S)nnc3-c3cccn3)CC2)CC1c1ncc[nH]1</chem>	-7.5
12	72 (H bond)	<chem>C=CCNC(=O)C(C#N)=c1sc(=CC=C2CCC(C)(C)C2)c(=O)n1CCCC(C)(C)C</chem>	-7.4
13	62	<chem>CCC=CCOc1cccc1-c1cccc2ccccc12</chem>	-7.3
14	4 (H bond)	<chem>COCC(O)CN1CCC2=C(c3ccccc3)CCC(=O)C2C1</chem>	-7.2
15	39	<chem>CCN(CCN1CCN(C(=O)CCCc2nnc(-c3ccccc3)o2)CC1)Cc1cccn1</chem>	-7.2
16	42	<chem>C=CCCCCN(C)C(=O)CC12CC3CC(C1)CC(C(F)(F)F)(C3)C2c1cccn1</chem>	-7.2
17	66 (H bond)	<chem>CC1CN(C(=O)OC(C)(C)C)C(C)CC1CCNCCc1c[nH]nc1C(C)(C)C</chem>	-7.2
18	50	<chem>CC#CC(=O)N1CCc2ccc(-c3ccccc3)cc2C1</chem>	-7.1
19	64 (H bond)	<chem>CCCC(CC1CCCC1)C(=O)N(C)c1ccc(CNC(=O)C2CC2C(C)(C)C)o1</chem>	-7.1
20	15	<chem>CCOCC(=O)NCCn1c2ccccc2c2ccccc21</chem>	-7.0

**Figure 16.** 2D representation of the two purchasable initial ligands found in ZINC. (A) Ligand 18 (ZINC ID: ZINC1160126598) (B) ligand 20 (ZINC ID: ZINC345426987).

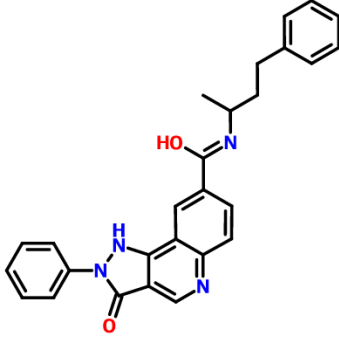
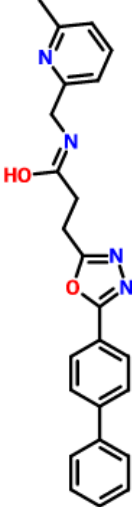
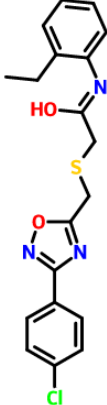
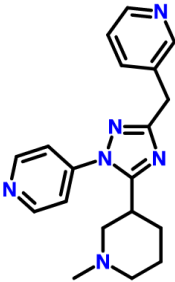
Similar ligand search provided also results. Table 5 shows the similarity percentage of every similar ligand generated from the ALDRICH MARKET, in correlation with the initial 20 ligands. Every similar ligand is symbolized as “X-S” where “X” represents the initial ligand and “-S” the fact that it is similar. All the similar ligands have a similarity score above or equal to 70%, with 33-S being the only exception.

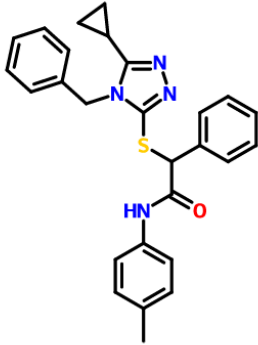
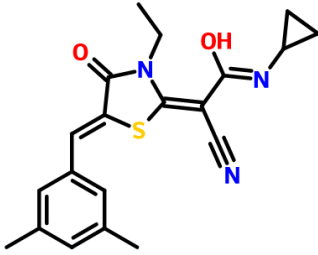
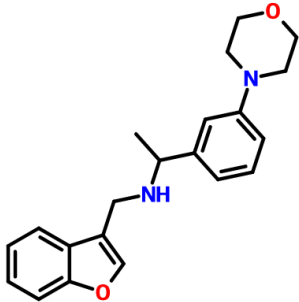
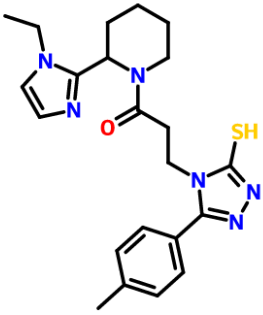
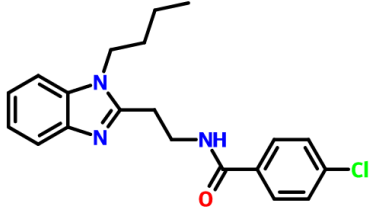
**Table 5.** Similar marketed ligands to the initial ones

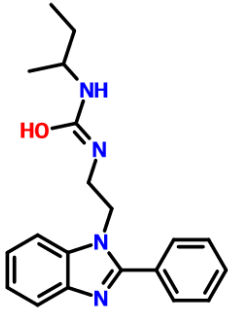
<b>X-S</b>	<b>Ligand in Initial Table</b>	<b>Top Similarity (%) in ALDRICH MARKET SELECT</b>	<b>Structure Id</b>
18-S	18 (H bond)	78	36546380
27-S	27	84	31652407
78-S	78	68	508624152
26-S	26 (H bond)	83	111261804
33-S	33	63	363722168
81-S	81	82	979600813
12-S	12 (H bond)	70	723643644
28-S	28	85	514732986
36-S	36	97	39428849
45-S	45	76	111239218
73-S	73	78	448958304
72-S	72 (H bond)	75	288519203
62-S	62	80	129708610
4-S	4 (H bond)	68	369250970
39-S	39	82	181956582
42-S	42	73	665275868
66-S	66 (H bond)	86	388600896
50-S	50	80	363922686
64-S	64 (H bond)	71	1060926810
15-S	15	88	29227343

After docking the abovementioned ligands individually, the final pooled ligands with a docking score exceeding -7 kcal/mol were 10. These 10 ligands are presented in Table 6 below.

**Table 6.** Docking scores & 2D representation of the similar marketed ligands.

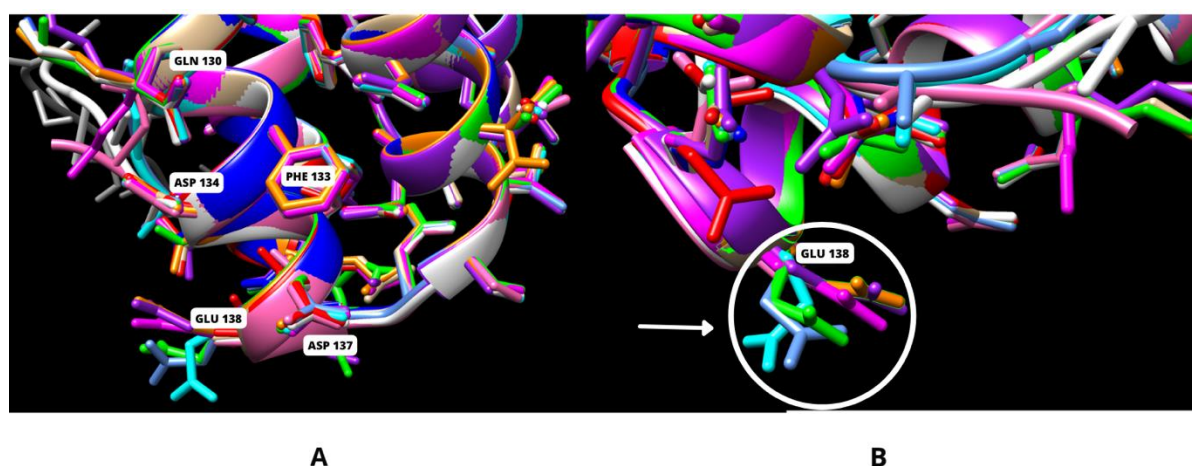
Similar marketed ligands	Descending Docking score	2D Representation
33-S	-8.6	
39-S	-8.3	
45-S	-8.0	
81-S	-7.9	

18-S	-7.8	
72-S	-7.7	
12-S	-7.6	
73-S	-7.5	
27-S	-7.3	

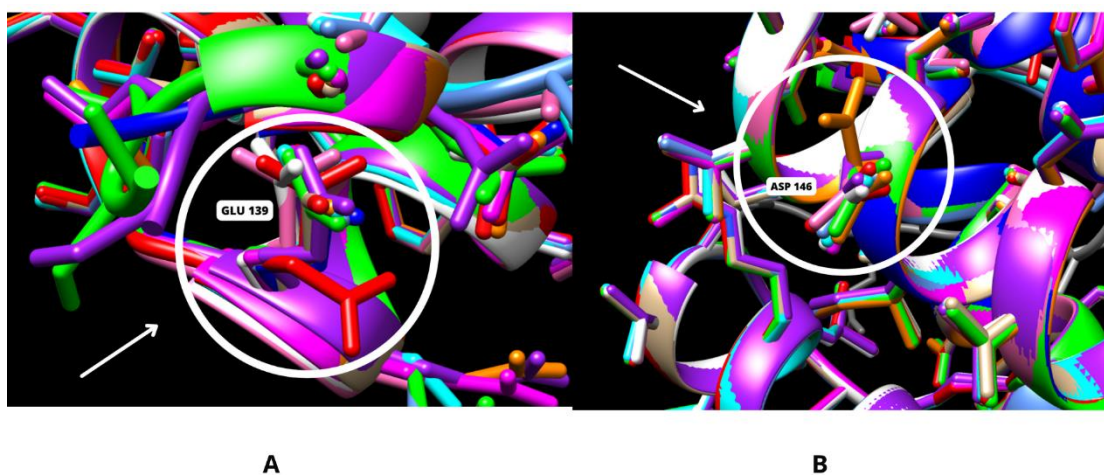
26-S	-7.3	
<p>*2D representation is a result of the conversion of the InChi of each X-S ligand with Cactus Chemical Identifier Resolver (<a href="https://cactus.nci.nih.gov/chemical/structure">https://cactus.nci.nih.gov/chemical/structure</a> accessed April 2022)</p>		

Docking of every single of the abovementioned ligands, both X and X-S, concluded in a specific amino acid trend. In all cases of pi-pi interactions, Phe133 was the responsible residue. Moreover, in the presence of H bonds, Asp and Asn were the only residues forming these bonds. Furthermore, the majority of the residues within a radius of 5 Å are Asn and Asp, with Asp being present 70 times and Asn 56 times respectively, among all the 20 X-S ligands bound in the N-terminal pore. In addition, Gly, Phe, and Ser are also present within this specific radius.

Regarding the superimposition of each monomer, the results showed some alternations in the side chains of specific amino acids, which are implicated in the binding of the ligand. These alternations are depicted below in Fig. 17 and 18.



**Figure 17.** Visual representation of the superimposed monomer side chains. Gln130, Phe133, Asp134, and Asp137 show slight to no change (A). Glu138 side chains appear to have different orientations (B).



**Figure 18.** Visual representation of the superimposed monomer side chains. Glu139 appears to have slightly different orientations among the side chains of the different monomers (A). Asp146 side chains appear to have slightly different orientations among the side chains of the different monomers (B).



## 5 Discussion

A new cost effective and faster purification scheme for Dpr by keeping the hexahistidine affinity tag was carried out. Ten crystallization conditions were identified, suggesting that the hexahistidine tag can be kept intact during crystallization. Out of these 10 conditions, optimization was performed in 3 of them, leaving an open area for further optimization of the 7 remaining ones, if required. Keeping the tag from the initiation of the protein production all the way until the crystallization, will minimize the time used for purification, decrease the cost of materials since no thrombin for cleavage is needed, and eventually make the whole process simpler. However, the protein showed aggregation problems and precipitated during storage. This is something that needs further investigation. It is possible that the pH used for protein storage (pH 7.0) may need to change. As the protein purified with HisPur™ Ni-NTA affinity column at pH 8, higher pH of the storage buffer may be better to keep the protein in soluble form. More studies need to be performed in order to optimize first the solubility of the protein. Protein aggregation might lead to non-equal protein deposition during the screening procedure, leading to irreproducible results due to the difference in protein amount and homogeneity.

Out of the 10 conditions that produced crystals, three were selected to provide crystals for data collection in a synchrotron facility. The majority of the conditions would need to be optimized by varying the concentration of the precipitant (PEG) and the pH, something that did not happen initially due to the time restriction of this particular project. Crystals of condition 7 were visually bigger than the ones of conditions 9 and 10, however, the crystals did not diffract in the X-ray beam yet no data were collected, highlighting once again the fact that the appearance of a crystal is not indicative of its diffraction properties. The next step would be to start optimizing condition 9 which provided the crystal that gave the acquired diffraction pattern as depicted in Figure 13. Having already received a good diffraction pattern, it will be our first approach toward collecting an even higher resolution data set. Then, optimization of condition 7 would follow. It is important to understand that each protein is unique, and this is also translated into the fact that the crystallization conditions are unique for each protein or protein complex (le Maire et al., 2020). There is neither a silver lining nor a golden rule regarding how to accomplish effortless crystallization conditions for proteins, especially for proteins carrying a histidine tag. Data are divisive, suggesting that the tag either helps or

prevents crystallization, depending for example on the protein, the flexibility of the tag, and the position of the tag at either the N- or the C-terminal (Malhotra, 2009).

Regarding the computational part, this study represents the first approach for ligandability studies and ligand library creation, so no other references exist to compare the results. As depicted from the outcome, the N-terminal pores are indeed competent to get blocked with ligands. The 10 marketed ligands could provide the starting point for further exploration and assessment of the *in-silico* method in the laboratory. Co-crystallization is a valuable step hoping to provide a structure that can be further utilized computationally. Moreover, biochemical assays could be used for ligand action assessment like those used previously (Kauko et al., 2006). In addition, the implication of each ligand within specific pathways needs to be examined. Thus, heatmaps need to be created for these ligands showing their interaction within an individual's body and highlighting drug-drug interactions. Aromaticity of all the marketed ligands is the first common feature that groups the ligands together, a characteristic that still has not been translated to being beneficial, random, or harmful for the final aim of creating a drug regimen. Further research needs to be conducted in order to fully understand the interactions occurring in the N-terminal pores.

Another approach for *ssDpr* would be to collect data with evolutionary close Dps-like proteins, as proposed by the phylogenetic tree of Dps-family representatives in the review of Haikarainen & Papageorgiou, 2009. Structural data and computational results of docking (if available) could be utilized for a better comprehension of our results in the structural part.

## 6 Conclusions

The main conclusions of this thesis are as follows:

- 6xHis-tag can be left intact in both the purification and crystallization procedures of *ssDpr*.
- 6xHis-tagged *ssDpr* can be crystallized in various conditions.
- The 6xHis-tagged *ssDpr* appears to precipitate easily during storage and pH may affect its solubility.
- *ssDpr* ligandability is existent with both *de novo* and marketed ligands that are able to occupy the N-terminal pores.
- Ligands with aromatic rings could be involved in pi-pi interactions with Phe133.

## 7 Abbreviations

<b>Å</b>	Angstrom ( $10^{-10}$ m)
<b>Au</b>	Gold
<b>CT</b>	Computed Tomography
<b>Dpr</b>	Dps-like peroxide resistance protein
<b>Dps</b>	DNA-binding protein from starved cells
<b>Fe</b>	Iron
<b>HP-NAP</b>	Neutrophil-activating protein from <i>Helicobacter pylori</i>
<b>MRI</b>	Magnetic Resonance Imaging
<b>MW</b>	Molecular weight
<b>RT</b>	Room temperature
<b>ROS</b>	Reactive oxygen species
<b>XAS</b>	X-ray absorption spectroscopy
<b>XRC</b>	X-ray crystallography

## References

- Almiron, M., Link, A. J., Furlong, D., & Kolter, R. (1992). A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes and Development*, 6(12 B), 2646–2654. <https://doi.org/10.1101/gad.6.12b.2646>
- Andrews, S. C. (2010). The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochimica et Biophysica Acta - General Subjects*, 1800(8), 691–705. <https://doi.org/10.1016/j.bbagen.2010.05.010>
- Aumiller, W., Uchida, M., & Douglas, T. (2018). Protein cage assembly across multiple length scales. *Chemical Society Reviews*, 176(3), 139–148. <https://doi.org/10.1039/c7cs00818j>.Protein
- Avershina, E., Shapovalova, V., & Shipulin, G. (2021). Fighting Antibiotic Resistance in Hospital-Acquired Infections: Current State and Emerging Technologies in Disease Prevention, Diagnostics and Therapy. *Frontiers in Microbiology*, 12(July). <https://doi.org/10.3389/fmicb.2021.707330>
- Bozzi, M., Mignogna, G., Stefanini, S., Barra, D., Longhi, C., Valenti, P., & Chiancone, E. (1997). A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*. *Journal of Biological Chemistry*, 272(6), 3259–3265. <https://doi.org/10.1074/jbc.272.6.3259>
- Cheng, Q. Di, Chen, R. Q., He, J., Li, D. W., Yang, F., Liu, Y. M., Lu, Q. Q., Dong, C., & Yin, D. C. (2017). Effect of the weather conditions during solution preparation on lysozyme crystallization. *Journal of Applied Crystallography*, 50(5), 1341–1351. <https://doi.org/10.1107/S1600576717011086>
- Cho, S. J., Idrobo, J. C., Olamit, J., Liu, K., Browning, N. D., & Kauzlarich, S. M. (2005). Growth mechanisms and oxidation resistance of gold-coated iron nanoparticles. *Chemistry of Materials*, 17(12), 3181–3186. <https://doi.org/10.1021/cm0500713>
- Chou, S., Shau, Y., Wu, P., & Yang, Y. (2010). *Chen. JACS. 2010, 132, 13270-13278. In Vitro and in Vivo Studies of FePt Nanoparticles for Dual Modal CT-MRI Molecular Imaging.pdf. 14, 13270–13278.*
- Dadinova, L. A., Chesnokov, Y. M., Kamyshinsky, R. A., Orlov, I. A., Petoukhov, M. V., Mozhaev, A. A., Soshinskaya, E. Y., Lazarev, V. N., Manuvera, V. A., Orekhov, A. S., Vasiliev, A. L., & Shtykova, E. V. (2019). Protective Dps–DNA co-crystallization in stressed cells: an in vitro structural study by small-angle X-ray scattering and cryo-electron tomography. *FEBS Letters*, 593(12), 1360–1371.

3468.13439

- Daniel, E., Maksimainen, M. M., Smith, N., Ratas, V., Biterova, E., Murthy, S. N., Tanvir Rahman, M., Kiema, T. R., Sridhar, S., Cordara, G., Dalwani, S., Venkatesan, R., Prilusky, J., Dym, O., Lehtio, L., Kristian Koski, M., Ashton, A. W., Sussman, J. L., & Wierenga, R. K. (2021). IceBear: An intuitive and versatile web application for research-data tracking from crystallization experiment to PDB deposition. *Acta Crystallographica Section D: Structural Biology*, *77*, 151–163. <https://doi.org/10.1107/S2059798320015223>
- Franceschini, S., Ceci, P., Alaleona, F., Chiancone, E., & Ilari, A. (2006). Antioxidant Dps protein from the thermophilic cyanobacterium *Thermosynechococcus elongatus*: An intrinsically stable cage-like structure endowed with enhanced stability. *FEBS Journal*, *273*(21), 4913–4928. <https://doi.org/10.1111/j.1742-4658.2006.05490.x>
- Frenkiel-Krispin, D., Levin-Zaidman, S., Shimoni, E., Wolf, S. G., Wachtel, E. J., Arad, T., Finkel, S. E., Kolter, R., & Minsky, A. (2001). Regulated phase transitions of bacterial chromatin: A non-enzymatic pathway for generic DNA protection. *EMBO Journal*, *20*(5), 1184–1191. <https://doi.org/10.1093/emboj/20.5.1184>
- Frey, P. A., & Reed, G. H. (2012). The ubiquity of iron. *ACS Chemical Biology*, *7*(9), 1477–1481. <https://doi.org/10.1021/cb300323q>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144. <https://doi.org/10.1145/3422622>
- Grant, R. A., Filman, D. J., Finkel, S. E., Kolter, R., & Hogle, J. M. (1998). The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nature Structural Biology*, *5*(4), 294–303. <https://doi.org/10.1038/nsb0498-294>
- Gupta, A. K., & Gupta, M. (2005). Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. *Biomaterials*, *26*(18), 3995–4021. <https://doi.org/10.1016/j.biomaterials.2004.10.012>
- Haikarainen, T., & Papageorgiou, A. C. (2009). Dps-like proteins: Structural and functional insights into a versatile protein family. *Cellular and Molecular Life Sciences*, *67*(3), 341–351. <https://doi.org/10.1007/s00018-009-0168-2>
- Hébraud, M., & Guzzo, J. (2000). The main cold shock protein of *Listeria monocytogenes* belongs to the family of ferritin-like proteins. *FEMS Microbiology Letters*, *190*(1), 29–34. [https://doi.org/10.1016/S0378-1097\(00\)00310-4](https://doi.org/10.1016/S0378-1097(00)00310-4)
- Herrmann, I., Urner, M., Hasler, M., Roth-Z'Graggen, B., Aemisegger, C., & Baulig, W.

- (2011). Iron core/shell nanoparticles as magnetic drug carriers: possible interactions with the vascular compartment. *Nanomedicine*, 6(7), 1199–1213.
- Imlay, J. (2008). Cellular Defenses against Superoxide and Hydrogen Peroxide. *Annual Review Of Biochemistry*, 34(1), 47–50.  
<https://doi.org/10.1146/annurev.biochem.77.061606.161055>.Cellular
- Ito, A., Honda, H., & Kobayashi, T. (2006). Cancer immunotherapy based on intracellular hyperthermia using magnetite nanoparticles: A novel concept of “heat-controlled necrosis” with heat shock protein expression. *Cancer Immunology, Immunotherapy*, 55(3), 320–328. <https://doi.org/10.1007/s00262-005-0049-y>
- Jiménez, J., Škalič, M., Martínez-Rosell, G., & De Fabritiis, G. (2018). KDEEP: Protein-Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. *Journal of Chemical Information and Modeling*, 58(2), 287–296.  
<https://doi.org/10.1021/acs.jcim.7b00650>
- Karas, V. O., Westerlaken, I., & Meyer, A. S. (2015). The DNA-binding protein from starved cells (Dps) utilizes dual functions to defend cells against multiple stresses. *Journal of Bacteriology*, 197(19), 3206–3215. <https://doi.org/10.1128/JB.00475-15>
- Kauko, A., Haataja, S., Pulliainen, A. T., Finne, J., & Papageorgiou, A. C. (2004). Crystal structure of Streptococcus suis Dps-like peroxide resistance protein Dpr: Implications for iron incorporation. *Journal of Molecular Biology*, 338(3), 547–558.  
<https://doi.org/10.1016/j.jmb.2004.03.009>
- Kauko, A., Pulliainen, A. T., Haataja, S., Meyer-Klaucke, W., Finne, J., & Papageorgiou, A. C. (2006). Iron Incorporation in Streptococcus suis Dps-like Peroxide Resistance Protein Dpr Requires Mobility in the Ferroxidase Center and Leads to the Formation of a Ferrihydrite-like Core. *Journal of Molecular Biology*, 364(1), 97–109.  
<https://doi.org/10.1016/j.jmb.2006.08.061>
- le Maire, A., Germain, P., & Bourguet, W. (2020). Protein-protein interactions in the regulation of RAR–RXR heterodimers transcriptional activity. In *Methods in Enzymology* (1st ed., Vol. 637). Elsevier Inc. <https://doi.org/10.1016/bs.mie.2020.02.007>
- Lemire, J., Alhasawi, A., Appanna, V. P., Tharmalingam, S., & Appanna, V. D. (2017). Metabolic defence against oxidative stress: the road less travelled so far. *Journal of Applied Microbiology*, 123(4), 798–809. <https://doi.org/10.1111/jam.13509>
- Liu, Y., Grimm, M., Dai, W. tao, Hou, M. chun, Xiao, Z. X., & Cao, Y. (2020). CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica*, 41(1), 138–144. <https://doi.org/10.1038/s41401-019-0228-6>

- Lunha, K., Chumpol, W., Samngamnim, S., Jiemsup, S., Assavacheep, P., & Yongkiettrakul, S. (2022). Antimicrobial Susceptibility of *Streptococcus suis* Isolated from Diseased Pigs in Thailand, 2018–2020. *Antibiotics*, *11*(3), 410.  
<https://doi.org/10.3390/antibiotics11030410>
- Lushchak, V. I. (2001). Oxidative stress and mechanisms of protection against it in bacteria. *Biokhimiya*, *66*(5), 592–609.
- Macedo, S., Romão, C. V., Mitchell, E., Matias, P. M., Liu, M. Y., Xavier, A. V., LeGall, J., Teixeira, M., Lindley, P., & Carrondo, M. A. (2003). The nature of the di-iron site in the bacterio-ferritin from *Desulfovibrio desulfuricans*. *Nature Structural Biology*, *10*(4), 285–290. <https://doi.org/10.1038/nsb909>
- Malhotra, A. (2009). Chapter 16 Tagging for Protein Expression. In *Methods in Enzymology* (1st ed., Vol. 463, Issue C). Elsevier Inc. [https://doi.org/10.1016/S0076-6879\(09\)63016-0](https://doi.org/10.1016/S0076-6879(09)63016-0)
- Miller, J. L. (2013). Iron deficiency anemia: A common and curable disease. *Cold Spring Harbor Perspectives in Medicine*, *3*(7). <https://doi.org/10.1101/cshperspect.a011866>
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., ... Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, *399*(10325), 629–655.  
[https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- Nicodème, M., Perrin, C., Hols, P., Bracquart, P., & Gaillard, J. L. (2004). Identification of an Iron-Binding Protein of the Dps Family Expressed by *Streptococcus thermophilus*. *Current Microbiology*, *48*(1), 51–56. <https://doi.org/10.1007/s00284-003-4116-3>
- Palmieri, C., Varaldo, P. E., & Facinelli, B. (2011). *Streptococcus suis*, an emerging drug-resistant animal and human pathogen. *Frontiers in Microbiology*, *2*(NOV), 1–6.  
<https://doi.org/10.3389/fmicb.2011.00235>
- Peña, M. M. O., & Bullerjahn, G. S. (1995). The DpsA Protein of *Synechococcus* sp. Strain PCC7942 Is a DNA-binding Hemoprotein. *Journal of Biological Chemistry*, *270*(38), 22478–22482. <https://doi.org/10.1074/jbc.270.38.22478>
- Perrin, C., Guimont, C., Bracquart, P., & Gaillard, J. L. (1999). Expression of a new cold shock protein of 21.5 kDa and of the major cold shock protein by *Streptococcus thermophilus* after cold shock. *Current Microbiology*, *39*(6), 342–347.  
<https://doi.org/10.1007/s002849900469>



- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612.  
<https://doi.org/10.1002/jcc.20084>
- Shieh, D. Bin, Cheng, F. Y., Su, C. H., Yeh, C. S., Wu, M. T., Wu, Y. N., Tsai, C. Y., Wu, C. L., Chen, D. H., & Chou, C. H. (2005). Aqueous dispersions of magnetite nanoparticles with NH<sub>3</sub><sup>+</sup> surfaces for magnetic manipulations of biomolecules and MRI contrast agents. *Biomaterials*, *26*(34), 7183–7191.  
<https://doi.org/10.1016/j.biomaterials.2005.05.020>
- Singh, N., Jenkins, G. J. S., Asadi, R., & Doak, S. H. (2010). Potential toxicity of superparamagnetic iron oxide nanoparticles (SPION). *Nano Reviews*, *1*(1), 5358.  
<https://doi.org/10.3402/nano.v1i0.5358>
- Skalic, M., Jiménez, J., Sabbadin, D., & De Fabritiis, G. (2019). Shape-Based Generative Modeling for de Novo Drug Design. *Journal of Chemical Information and Modeling*, *59*(3), 1205–1214. <https://doi.org/10.1021/acs.jcim.8b00706>
- Stillman, T. J., Upadhyay, M., Norte, V. A., Sedelnikova, S. E., Carradus, M., Tzokov, S., Bullough, P. A., Shearman, C. A., Gasson, M. J., Williams, C. H., Artymiuk, P. J., & Green, J. (2005). The crystal structures of *Lactococcus lactis* MG1363 Dps proteins reveal the presence of an N-terminal helix that is required for DNA binding. *Molecular Microbiology*, *57*(4), 1101–1112. <https://doi.org/10.1111/j.1365-2958.2005.04757.x>
- Storz, G., & Imlay, J. A. (1999). Oxidative stress. *Current Opinion in Microbiology*, *2*(2), 188–194.
- Surgutskaya, N. S., Trusova, M. E., Slepchenko, G. B., Minin, A. S., Pershina, A. G., Uimin, M. A., Yermakov, A. E., & Postnikov, P. S. (2017). Iron-core/carbon-shell nanoparticles with intrinsic peroxidase activity: New platform for mimetic glucose detection. *Analytical Methods*, *9*(16), 2433–2439. <https://doi.org/10.1039/c7ay00598a>
- Vouga, M., & Greub, G. (2016). Emerging bacterial pathogens: The past and beyond. *Clinical Microbiology and Infection*, *22*(1), 12–21. <https://doi.org/10.1016/j.cmi.2015.10.010>
- Williams, R. J. P. (1982). Free manganese(II) and iron(II) cations can act as intracellular cell controls. *FEBS Letters*, *140*(1), 3–10. [https://doi.org/10.1016/0014-5793\(82\)80508-5](https://doi.org/10.1016/0014-5793(82)80508-5)
- Williams, S. M., Chandran, A. V., Prakash, S., Vijayan, M., & Chatterji, D. (2017). A Mutation Directs the Structural Switch of DNA Binding Proteins under Starvation to a Ferritin-like Protein Cage. *Structure*, *25*(9), 1449–1454.e3.  
<https://doi.org/10.1016/j.str.2017.07.006>

- Wu, P. C., Wang, W. S., Huang, Y. T., Sheu, H. S., Lo, Y. W., Tsai, T. L., Shieh, D. Bin, & Yeh, C. S. (2007). Porous iron oxide based nanorods developed as delivery nanocapsules. *Chemistry - A European Journal*, *13*(14), 3878–3885.  
<https://doi.org/10.1002/chem.200601372>
- Wu, Y. N., Shieh, D. Bin, Yang, L. X., Sheu, H. S., Zheng, R., Thordarson, P., Chen, D. H., & Braet, F. (2018). Characterization of iron core-gold shell nanoparticles for anti-cancer treatments: Chemical and structural transformations during storage and use. *Materials*, *11*(12), 1–14. <https://doi.org/10.3390/ma11122572>
- Yamamoto, Y., Poole, L. B., Hantgan, R. R., & Kamio, Y. (2002). An iron-binding protein, Dpr, from *Streptococcus mutans* prevents iron-dependent hydroxyl radical formation in vitro. *Journal of Bacteriology*, *184*(11), 2931–2939.  
<https://doi.org/10.1128/JB.184.11.2931-2939.2002>
- Zanotti, G., Papinutto, E., Dundon, W. G., Battistutta, R., Seveso, M., Giudice, G. Del, Rappuoli, R., & Montecucco, C. (2002). Structure of the neutrophil-activating protein from *Helicobacter pylori*. *Journal of Molecular Biology*, *323*(1), 125–130.  
[https://doi.org/10.1016/S0022-2836\(02\)00879-3](https://doi.org/10.1016/S0022-2836(02)00879-3)
- Zhou, W. L., Carpenter, E. E., Lin, L., Kumbhar, A., Sims, J., & O'Connor, C. J. (2001). Nanostructures of gold coated iron core-shell nanoparticles and the nanobands assembled under magnetic field. *European Physical Journal D*, *8*(3), 289–292.  
<https://doi.org/10.1007/s100530170112>
- Zhu, J. Y., Zhang, R., Pathak, D., Darrell, T., Efros, A. A., Wang, O., & Shechtman, E. (2017). Toward multimodal image-to-image translation. *Advances in Neural Information Processing Systems*, *2017-Decem*(1), 466–477.

## 8 Appendices

### 8.1 Appendix 1. Master's Thesis Supervision Plan



#### Appendix 1: Master's Thesis Supervision Plan

#### Master's Degree Programme in Biomedical Sciences, Drug Discovery and Development

##### Student

Last name Tatsis	First name(s) Polychronis
Student number 524409	Email address potats@utu.fi
The working title of the MSc Thesis The iron-entry pores in Dps-like proteins as potential drug-design target: Structural investigations	

##### Responsible Supervisor

Last name Papageorgiou	First name(s) Anastassios	Title/Degree Adjunct Professor, PhD
Affiliation Turku Bioscience, University of Turku		Email address anapap@utu.fi

##### Second Supervisor (optional)

Last name Poudel	First name(s) Nirmal	Title/Degree Doctoral candidate
Affiliation University of Turku		Email address nirpou@utu.fi

##### Third Supervisor (only if neither Responsible or Second Supervisor is affiliated at UTU)

Last name	First name(s)	Title/Degree
Affiliation		Email address

Thesis supervision involves regular discussions on the progress of the thesis between the Supervisor(s) and the Student. In good supervision, both the Supervisor(s) and the Student must have a clear view of their own role in the Thesis process and a shared understanding about the temporal progress of the Thesis. In order for the timetable of the thesis work to be followed, time limits have to be agreed upon with the Supervisor(s) also for the different stages and sections of the Thesis. The purpose of good supervision is to support the Student working on the Thesis with both personal as well as group guidance. Good supervision provides the Student with a safe and guided learning environment for carrying out the Thesis independently.

Both the Supervisor(s) and the Student have their rights, responsibilities and duties in the supervision relationship.

##### The Student:



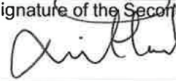
- Right to receive supervision during the thesis process. The rules of the supervision are agreed beforehand with the supervisor(s).
- Responsible for taking the agreed time limits for the progress of the thesis into account.
- Responsible for carrying out the thesis independently complying with the rules of ethical research.
- Responsible for the progress of the thesis and for taking initiative when in need of guidance and facing any possible problems.
- Responsible of having regular discussions with the supervisor about the progress of the thesis.
- Responsible for immediately informing the supervisor of all such changes which significantly affect the work in progress, the continuation of the work, or the schedule.
- Agrees that the Research Plan and MSc Thesis are checked for plagiarism.

**Master's Degree Programme in Biomedical Sciences, Drug Discovery and Development**

**The Supervisor(s):**

- Committed to guide the thesis work regularly. The rules of the guidance are agreed beforehand with the student.
- Committed to guide the thesis work according to the agreed timetable. The agreed timetable should not be exceeded.
- Supports the thesis work and gives critical and constructive feedback.
- Responsible to give advice and takes into consideration that the student is carrying out his or hers first actual research and many of the methods and requirements can be unfamiliar to the student. If the thesis is carried out in a research project, the Student and the Supervisor have to discuss beforehand what does it mean in practice for the Thesis.
- Committed to have regular discussions with the Student about the progression of the Thesis.
- Ensures the suitability of the work for MSc Thesis.
- Informs the student beforehand about the evaluation principles of the Thesis, as well as of the plagiarism check (Turnitin system) as a part of the process.

Both the Student and the Supervisor(s) are responsible for committing to follow the ethical guidelines of the research. Master's Thesis Supervision Plan comes into effect after it has been signed. Thesis work can begin only after the Supervision Plan and the Project Plan have been approved by the Formal Examiner.

Date 1-09-2021	Signature of the Student and clarification  POLYCHRONIS TATSIS
Date 1-09-2021	Signature of the Responsible Supervisor and clarification  A.C. PAPAGEORGIOU
Date 1-09-2021	Signature of the Second Supervisor and clarification  NIRMAL POUDEL
Date	Signature of the Third Supervisor and clarification
Date	Signature of the Formal Examiner and clarification

The signed and scanned form will be delivered to the Formal Examiner/Programme Coordinator via Moodle site.

## 8.2 Appendix 2. Mandatory mid-phase feedback



**UNIVERSITY  
OF TURKU**

Faculty of Medicine

Appendix 2: Mid-phase Feedback

### Master's Degree Programme in Biomedical Sciences, Drug Discovery and Development

Regular feedback on the performance of the Student during the MSc Thesis Project is very important for the learning experience. For the MSc Thesis Project, the Supervisor(s) should provide oral feedback and assess the progress together with the Student at the Mid-phase of the Thesis Project after the experimental phase is done.



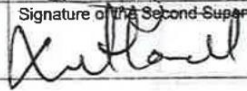
Both Supervisor(s) and Student sign the form, stating that this feedback session took place. The signed form is given to the Formal Reviewer or Programme Coordinator at the mandatory Mid-phase meeting of the course.

It is advisable to use the grading criteria of the MSc project work as a starting material for the feedback. The Mid-phase Feedback is also a good opportunity to discuss about the progress of the MSc Thesis Project and if some alternative approaches are needed.

The Student should also state clearly if there are some technical, time table-related or other problems that are preventing him/her to achieve the set goals.

#### Title of the MSc Thesis Project

The iron-entry pores in Dps-like proteins as potential drug-design target: Structural investigations

Date 14/11/2022	Signature of the Student and clarification  POLKCHRONIS TATSIS
Date 14/11/2022	Signature of the Responsible Supervisor and clarification  A. PAPAEFORGIOU
Date 14/11/2022	Signature of the Second Supervisor and clarification  NIRMAL POUDEL

The signed and scanned form will be delivered to the Formal Examiner/Programme Coordinator via Moodle site