Rafaela Carvalho Martins

Bachelor degree in Biochemistry

Functional and structural characterization of human GAPDH with a CPO-based probe

Dissertation for the Master degree in Biochemistry for Health

Supervisor: Dr. Marta C. Marques, iMM

Co-supervisor: Dr. Pedro M. Matias, ITQB-NOVA

November 2021

Rafaela Carvalho Martins

Bachelor degree in Biochemistry

Functional and structural characterization of human GAPDH with a CPO-based probe

Dissertation for the Master degree in Biochemistry for Health

Supervisor: Dr. Marta C. Marques, iMM

Co-supervisor: Dr. Pedro M. Matias, ITQB-NOVA

Júri

Presidente: Prof. Dra. Maria Teresa Nunes Mangas Catarino

Arguente: Prof. Dr. João B. Vicente

Vogal: Prof. Dra. Ana Maria de Jesus Bispo Varela Coelho

Instituto de Tecnologia Química e Biológica

November 2021

Direitos de Autor/Copyright

Acknowledgements

Before anything else, I would like to express my enormous gratitude to all the people who directly or indirectly contributed to this thesis project. Also, to the hosting entities and collaborators, ITQB-NOVA and iMM JLA for the excellent working conditions.

To my supervisor, Dr. Marta C. Marques, for not only trusting and respecting my work, but also for providing me a rich source of inspiration as a successful woman in science. I have a lot to thank you, not only on a professional level and for everything you taught me, but also on a personal level. You were my counselor and my listener. You had enormous patience and understanding beyond imagination with me. I can only thank you for everything, and thank you is not enough.

To my co-supervisor Dr. Pedro M. Matias, for letting me join this project and giving me the opportunity to work at the Macromolecular Crystallography Unit. For the vote of confidence and respect that enable me to always push forward and to grow as a scientist.

To Dr. Gonçalo Bernardes at GBernardes Lab (iMM JLA), for providing me with all the funds and tools necessary for me to complete my work. Also, to all my colleagues who welcomed me and made me feel part of the group throughout my journey, especially Bárbara Sousa and Maria Inês Albuquerque, for her truly and honest companionship, and support. I have to thank you for all the patience you had with me and for all the teaching you gave me. I am most thankful for meeting all of you.

To our collaborator at Cambridge University, Toby Journeaux, for making part of this project, for useful discussions and assistance with CPO-probe and cell culture work.

To my colleagues at the Macromolecular Crystallography Unit for the good working environment. In particular, to Diana Silva, for work support and companionship. Without your help and support none of this would be possible. Thank you for having accompanied me on this journey, for all your teachings and for your good disposition.

To my closest friends Rebeca Ramos and Maria Beatriz Felgueiras, I have a lot to thank you! Thanks for all the "you can do it", for all the good times when I just needed to calm down because things went wrong. Thanks for the nights spent in the lab with me just to keep me company. This work owes a lot to you! To Mariana Lima, Rui Matos e Mariana da Silva, for making my life away from home much, much easier. Thank all of you for always being by my side and making me believe it was possible when I didn't believe it myself. I just have to thank you for your enormous friendship!

À minha irmã Raquel, companheira da minha vida e a pessoa que melhor me conhece, e ao João, por me suportarem em todos os meus momentos de desespero, e aturarem o meu mau feitio quando tudo corria mal. Obrigada por serem o meu pilar, o meu porto de abrigo e o meu lugar seguro.

Por último, aos meus pais, Teresa e Fernando, que nunca deixaram de acreditar em mim e fizeram tudo para que conseguisse chegar onde estou. Agradeço-vos eternamente por tudo o que fizeram e continuam a fazer por mim. Devo-vos o meu sucesso, agora e sempre.

Table of Contents

Abstract

Cancer cells have a higher rate of glycolysis. Even in aerobic conditions, they exploit glycolysis as their main source of energy. The *Warburg effect* is the name for this phenomenon. As a result, efficient suppression of the glycolytic pathway is an important part of cancer treatment. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is an important enzyme that participates in numerous cellular functions and has been found to promote cancer growth and metastasis in several cancer types. As a result, inhibiting GAPDH contribution in glycolysis is critical for GAPDH functional research and cancer therapy yet, there have only been a few GAPDH inhibitors described.

Since crystallographic information is essential to understand the molecular basis of hGAPDH inhibition by covalent inhibitors, we establish an expression system bearing at the N-terminal region, a Strep-tag II fused to a TEV cleavage site using a typical expression strain, *E. coli* BL21(DE3), for the production of the recombinant human GAPDH. Here we report the full biochemical and biophysical characterization of human GAPDH alone and in complex with a monosubstituted cyclopropenone (CPO). CPOs have recently been shown to react selectively with N-terminal cysteine residues in proteins containing multiple cysteine residues. Evaluation of the protein sample and the hGAPDH-CPO complex quality was performed before crystallization trials, using SDS-PAGE, Western Blot, Dynamic Light Scattering, Circular Dichroism, Differential Scanning Fluorimetry and LC-MS. Moreover, we set up crystallization trials to generate crystals for X-ray diffraction analysis.

The X-ray crystal structure of hGAPDH in complex with NADH was determined to 2.11 Å resolution. hGAPDH crystallized in space group *I*222 with three independent molecules in asymmetric unit. Despite many attempts, the structure of hGAPDH with the CPO-probe has not been determined.

Hopefully, the determination of the 3D-dimensional structure of the hGAPDH in complex with this CPO-probe may represent a major advance in future cancer therapy.

Keywords: Strep-tag II, human recombinant GAPDH, monosubstituted cyclopropenones, biochemical and biophysical assays, X-ray crystallography structure determination

Resumo

As células cancerígenas são caracterizadas por uma maior taxa glicolítica. Mesmo em condições aeróbicas, utilizam a glicólise como principal fonte de energia. Este fenómeno é conhecido como efeito de *Warburg*. Consequentemente, a inibição da via glicolítica é uma parte importante para o tratamento do cancro. A gliceraldeído-3-fosfato desidrogenase (GAPDH) é uma enzima importante que participa em várias funções celulares e promove a formação de metástases em vários tipos de cancro. Como resultado, a inibição desta enzima na glicólise é fundamental para a terapia do cancro. Porém, até ao momento, existem poucos inibidores descritos para a GAPDH.

Uma vez que a informação cristalográfica é essencial para compreender a base molecular da inibição da GAPDH por inibidores covalentes desenvolvemos um sistema de expressão que contém na região N-terminal uma Strep-tag II ligada a um local de clivagem TEV, usando um sistema de expressão comum *E. coli* BL21 (DE3), para a produção GAPDH recombinante (hGAPDH). Aqui descrevemos a caracterização bioquímica e biofísica completa da GAPDH recombinante humana, assim como em complexo com uma ciclopropenona monossubstituída (CPO). Recentemente, foi demonstrado que as CPOs reagem seletivamente com resíduos de cisteína N-terminal em proteínas contendo múltiplos resíduos deste resíduo. Avaliámos a qualidade da amostra purificada e do complexo hGAPDH-CPO previamente à cristalização, utilizando SDS-PAGE, Western Blot, Dynamic Light Scattering, Circular Dichroism, Differential Scanning Fluorimetry e LC-MS. Posteriormente, foram efetuados ensaios de cristalização de modo a gerar cristais para a análise de difração de raios-X.

A estrutura cristalina obtida por raios-X da hGAPDH em complexo com NADH foi determinada com uma resolução de 2.11 Å. A hGAPDH cristalizou no grupo espacial *I*222 com três moléculas independentes na unidade assimétrica. Apesar de muitas tentativas, a estrutura de hGAPDH com a CPO não foi determinada.

Este conhecimento servirá como base para estudos que visam o desenvolvimento de moléculas terapêuticas, de modo a superar as limitações existentes nas terapias atuais.

Palavras-chave: Strep-tag II, GAPDH recombinante humana, ciclopropenona monossubstituída, ensaios bioquímicos e biofísicos, determinação de estrutura por cristalografia de raio-X

List of Figures

List of Tables

List of Abbreviations

BPG	1,3-Bisphosphoglycerate
DOPAL	3,4-Dihydroxyphenylacetaldehyde
ATP	Adenosine-Three-Phosphate
|F calc|	Calculated Structure-Factor Amplitude
CD	Circular Dichroism
Cys	Cysteine
DMF or MMF	Di- Or Monomethyl Fumarate
DSF	Differential Scanning Fluorimetry
DMSO	Dimethyl Sulfoxide
DMEM	Dulbecco's Modified Eagle Medium
DLS	Dynamic Light Scattering
θ	Ellipticity
E. coli	*Escherichia Coli*
EDTA	Ethylenediamine Tetraacetic Acid
FBS	Fetal Bovine Serum
FTSA	Fluorescence Thermal Shift Assay
Glu	Glutamate
G3P	Glyceraldehyde-3-Phosphate
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
His	Histidine
IB	Inclusion Bodies
IAA	Iodoacetate
IPTG	Isopropyl B-D-Thiogalactopyranoside
KA	Koningic Acid
LC-MS	Liquid Chromatography–Mass Spectrometry
LB	Luria-Bertani Medium
Lys	Lysine
Tm	Melting Temperature

Met	Methionine
UAAs	Modification Of Natural Or Unnatural Amino Acids
MR	Molecular Replacement
CPO	Monosubstituted Cyclopropenone
TEMED	N,N,Nn-Tetramethylethylenedi
NAD^+	Nicotinamide Adenosine Dinucleotide
NEAA	Non-Essential Amino Acids
|F obs|	Observed Structure-Factor Amplitude
Phe	Phenylalanine
PBS	Phosphate Buffered Saline
PAGE	Polyacrylamide Gel Electrophoresis
PEG	Polyethylene Glycol
PCR	Polymerase Chain Reaction
Pro	Proline
hGAPDH	Recombinant Human Glyceraldehyde-3-Phosphate Dehydrogenase
RRMS	Relapsing-Remitting Ms
r.m.s.d.	Root-Mean-Square Deviation
Ser	Serine
ΔTm	Shift In The Melting Temperature
SDS	Sodium Dodecyl Sulphate
|Fhkl|	Structure Factor Amplitude
TEV	Tobacco Etch Virus
Trp	Tryptophan
Tyr	Tyrosine

Chapter 1

1. Introduction

1.1. Cancer – Epidemiology and molecular basis

Cancer is a wide class of diseases that can begin in practically any organ or tissue of the body and spread to other organs when abnormal cells proliferate intensely, invade adjacent regions of the body, and/or move to other organs.[1] It is considered to be a biological ilness that is characterized by a complex interaction of genetic and enviromental variables that coordinate carcinogenesis.[2] Cancer is a leading cause of death worldwide, accounting for nearly 10 million fatalities in 2020.[1] Lung, prostate, colorectal, stomach and liver cancer are the most common types of cancer in men, while breast, colorectal, lung, cervical and thyroid cancer are the most common among women (Figure 1.1).[1,3]

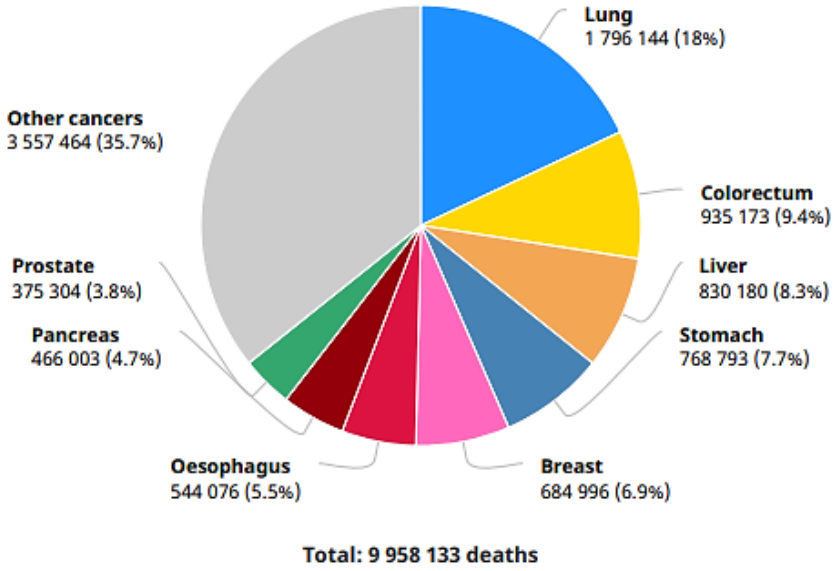

Figure 1.1. Number of deaths in 2020, both sexes, all ages. Data source: GLOBOCAN 2020, World Health Organization. (Adapted from [4]).

Cancer develops when normal cells are transformed into tumor cells in a multi-stage process that usually evolves from a pre-cancerous lesion to a malignant tumor. Physical carcinogens, such as ultraviolet and ionizing radiation; chemical carcinogens, such as asbestos and components of tobacco smoke and biological carcinogens, such as infections from certain viruses, bacteria, or parasites, can cause these changes.[1] Mutagenic events that modify a cell's genetic material have been shown to disrupt the pathways that control the cell's most basic functions.[2] The development of malignant conditions is caused by a build-up of numerous genetic abnormalities that result in the dysregulation of signaling pathways controlling cell proliferation, apoptosis, and DNA repair.[5] Cells with damaged DNA are generally eliminated by the body before

they become malignant. However, as we become older, our bodies' ability to do so decreases. This is one of the reasons why people are more likely to develop cancer later in life. Each person's cancer is made up of a unique set of genetic alterations and additional alterations will occur as the malignancy progresses. Different cells within the same tumor may have different genetic alterations.[6]

Proto-oncogenes, tumor suppressor genes, and DNA repair genes are all affected by the genetic alterations that contribute to cancer. These changes are commonly referred to as cancer "drivers". Proto-oncogenes play a role in normal cell division and proliferation. These genes can become cancer-causing genes (or oncogenes) if they are mutated in specific ways or are more active than usual, allowing cells to grow and survive when they shouldn't.[6] Tumor suppressor genes are also engaged in cell division and growth control. Certain mutations in tumor suppressor genes can cause cells to divide uncontrollably.[6] DNA repair genes are responsible for repairing damaged DNA. Cells with mutations in these genes are more likely to generate mutations in other genes and chromosome alterations, such as chromosome duplications and deletions. These alterations may lead the cells to become cancerous if they occur together.[6] Scientists have discovered that certain mutations are frequent in many types of cancer as they learn more about the molecular changes leading to cancer. Many anti-cancer medicines are now available that target cancer-causing gene alterations and these treatments are suitable for anyone with a cancer possessing the targeted mutation, regardless of where the tumor started developing.[6]

Globally, the incidence of cancer is increasing, placing enormous physical, emotional, and financial strain on individuals, families, communities, and health systems. Many health systems in low- and middle-income nations are unprepared to handle this burden, and many cancer patients around the world lack timely access to high-quality diagnosis and treatment. Many types of cancer survival rates are improving in countries with strong health systems, thanks to early detection, quality treatment, and survivorship care.[1,3]

1.2. Metabolic malignancy

Over the past few decades, the number of studies dealing with cancer have grown considerably, leading to a better understanding of this disease overall. Among the numerous discoveries made, some are worth mentioned here, namely that cancer is not a condition that can only be caused by genetic and epigenetic changes or that cancer cells are characterized by abnormal energy metabolism.[7] Indeed, cancer cells re-adjust their energy metabolism, an important step that helps them survive, evolve, and proliferate. Glucose absorption and lactic acid fermentation also increase as a result of this modification. Even in the presence of unimpaired mitochondrial integrity and aerobic conditions, glycolysis is a significant source of energy for cancer cells.[8]

1.2.1. Warburg Effect

The first steps towards unraveling cancer metabolism were taken more than 80 years ago, when Otto Warburg [9] discovered the so-called *Warburg effect*, in which cancer cells, both in anaerobic and aerobic environments, can generate adenosine triphosphate (ATP) that is biased toward lactic fermentation rather than mitochondrial oxidative phosphorylation (Figure 1.2).

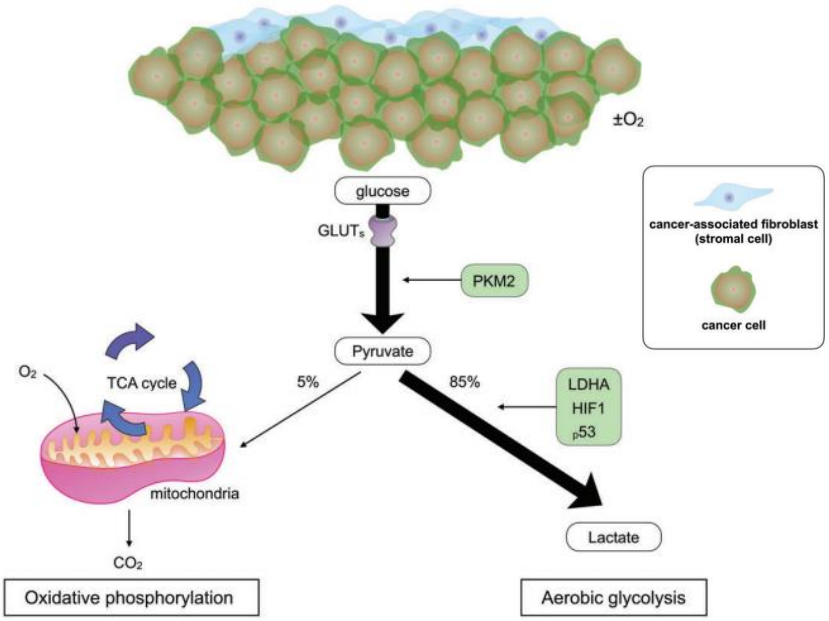

Figure 1.2. The *Warburg effect* is depicted as a schematic diagram. The *Warburg effect*, also known as aerobic glycolysis, turns glucose to lactate without causing oxidative phosphorylation in the mitochondria. It's regulated by proteins such as pyruvate kinase M2 (PKM2), lactate dehydrogenase A (LDHA), hypoxia-inducible factor-1, and p53. (Adapted from [7]).

In the presence of oxygen, the full oxidation of one glucose molecule (oxidative phosphorylation) in a cell creates 38 molecules of ATP, which is the primary cellular fuel.[10] The initial step of glucose cellular respiration (glycolysis, which occurs in the cytoplasm) generates only two ATP molecules and results in the production of two molecules of pyruvic acid. If a cell has access to oxygen, pyruvic acid is converted to acetyl-coenzyme A, which then enters the Krebs cycle (a citric acid cycle that takes place within mitochondria), followed by the electron transport chain process (which takes place on the inner mitochondrial membrane) that produces the majority of ATP molecules. On the other hand, if the cells are exposed to low oxygen levels, pyruvic acid is transformed into lactic acid rather than acetyl-coenzyme A, a process known as anaerobic cellular respiration (lactic acid cycle).[10] The net energy balance in the latter scenario is only two ATP molecules, making anaerobic glucose metabolism an extremely inefficient process energetically.[11]

While a deficit of oxygen at the cellular level can develop as a result of excessive physical activity (resulting in muscle soreness), it is also a hallmark of tumors that are highly invasive and rapidly proliferating.[10] As cancer cells expand, their rapid multiplication rate outpaces angiogenesis to the point where, while glucose supply to fast-growing cells may be sufficient, the absence of blood vessels prevents the full oxidation of glucose. Based on the differences between aerobic and anaerobic glucose metabolism, cancer cells (as well as healthy cells) in hypoxic conditions would require a 19-fold higher glucose intake to maintain the same metabolic state as well-oxygenated cells.[7,11] Cancer cells have a different energy profile than normal cells. They rely mostly on cytoplasmic glycolysis for energy generation, whereas normal cells rely more on oxidative phosphorylation and the Krebs cycle.[12] Lactate is generated aerobically in cancer cells and extruded, while no lactate is formed in normal cells undergoing aerobic glycolysis (Figure 1.3).[10,11]

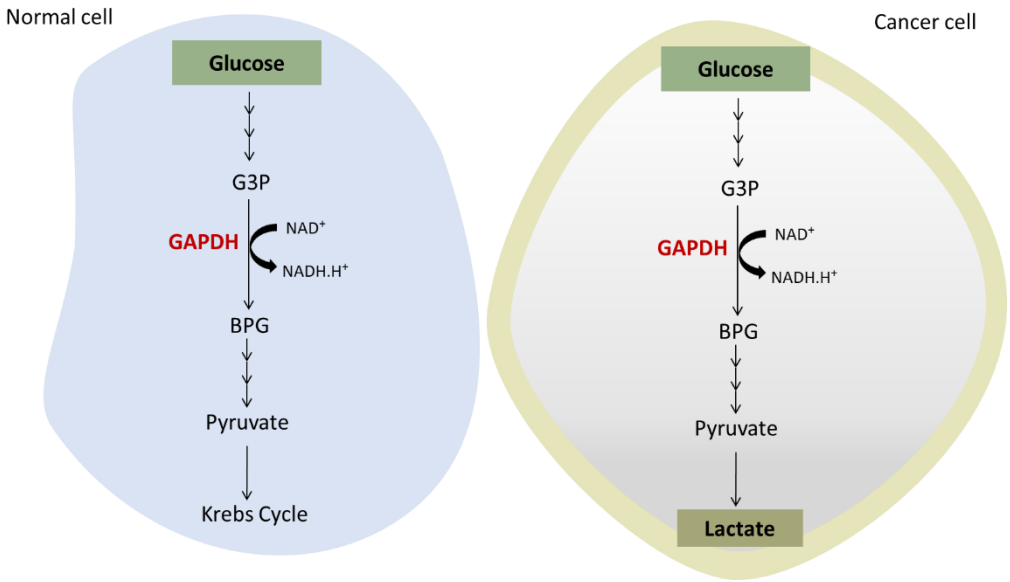

Figure 1.3. The dependency of cancer cells on glycolysis as a significant source of energy, even in the presence of oxygen, is one of the fundamental metabolic divergences between cancer cells and normal cells (*Warburg effect*). In the *Warburg effect*, glucose is catabolized to lactate, which is expelled into the microenvironment of cancer cells via monocarboxylate transporters, whereas glucose is metabolized to pyruvate, which is not expelled.

Even in the presence of oxygen, cancer cells exhibit glycolysis, lactate production, and mitochondrial respiration at the same time (*Warburg effect*).[13] Given that aerobic glycolysis produces significantly less ATP (only 2 moles of ATP per mole of glucose) than mitochondrial respiration (36 moles of ATP per mole of glucose), this mixed metabolic phenotype appears illogical.[14,15] For the preservation of this seemingly wasteful catabolic condition, several hypotheses have been presented. The *Warburg effect* appears to be a favorable catabolic condition for all rapidly growing mammalian cells with high glucose absorption capacity, according to Vasquez *et al.* in 2010.[14] It arises because, whereas aerobic glycolysis is less effective in terms of ATP output per glucose absorption than mitochondrial respiration, it is more efficient in terms of the needed solvent capacity.[14]

The complex nature of glycolysis, as well as the variations in its regulation between normal and cancer cells, offer methods for finding new anti-cancer targets and a wide range of options for developing highly selective anti-cancer drugs.[8] Because cancer cells require a lot of energy to survive, it's critical to understand energy metabolism based on the *Warburg effect* in order to design effective anti-cancer therapies that target cancer cells' energy generation route.

1.2.2. Tumorigenic glycolysis

As described in the previous section, cancer cells are characterized by an abnormal energy metabolism. In contrast to normal differentiated cells, which predominately depend on mitochondrial oxidative phosphorylation to generate energy, cancer cells rely on aerobic glycolysis, a phenomenon coined *Warburg effect*.[9] Cancer cells can survive, thrive, and proliferate in hypoxic conditions, which are frequent in tumors, thanks to a metabolic adaptation that includes rapid glucose intake coupled with lactate output. This overdependence on glycolysis, on the other hand, provides a chance to target cancer cells specifically.[16]

High glycolytic rates confer selective advantage by promoting uncontrolled growth and resistance to chemotherapy, so Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), a glycolytic key-enzyme will be discussed in further detail.

1.3. Glyceraldehyde-3-phosphate dehydrogenase and glycolysis

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a tetrameric enzyme that has long been known to have a role in energy metabolism and the synthesis of ATP and pyruvate in the cytoplasm via anaerobic glycolysis.[17]

Once its function was believed to have been unveiled to the last detail, GAPDH was thrown in the “housekeeping gene” list, and it was dismissed as having no substantial research influence. Recent research has discovered that GAPDH can perform a variety of tasks that are unrelated to its role in glycolysis, which include membrane fusion, cytoskeleton dynamics, DNA repair and RNA export.[18,19] These functions are thought to be regulated, at least in part, by post-translational modifications, oligomerization and subcellular localization, according to Tristan *et* al.[18] GAPDH has been discovered to be a sensor of intracellular and extracellular cell stress, capable of triggering either recovery or cell death signaling pathways.[18,20] Under normal physiological settings, active GAPDH stimulates the transition from anaerobic to pentose phosphate respiration in the cytoplasm. This enzyme, however, can be reversibly inactivated by S-thiolation of its sulfhydryl groups under oxidative stress, allowing cells to flip between metabolic tasks and maintain a balanced oxidation/reduction state.[21] Increased levels of GAPDH in the mitochondria, on the other hand, can trigger apoptotic processes, resulting in mitochondrial membrane permeabilization after the release of apoptosis-inducing factor and cytochrome c.[22] These new

numerous GAPDH functions, particularly those involving cell death mechanisms[20], may imply that this enzyme is involved in some illnesses when produced at abnormal levels.[18] There is some evidence to back this up, especially in the case of neurodegenerative diseases. GAPDH levels are observed to be higher in the post-mortem brains of patients with Parkinson's disease and Alzheimer's disease[23,24] and it has been tentatively suggested that GAPDH can interact with mutant huntingtin and β-amyloid peptides, respectively.[25,26]

However, its primary role is catalyzing the sixth step of glycolysis. In the presence of nicotinamide adenosine dinucleotide (NAD^+) and inorganic phosphate, GAPDH reversibly catalyzes the first step of the pathway, in which glyceraldehyde-3-phosphate (G3P) is converted into a high-energy phosphate compound 1,3-bisphosphoglycerate (BPG), as described in Figure 1.4.[27,28]

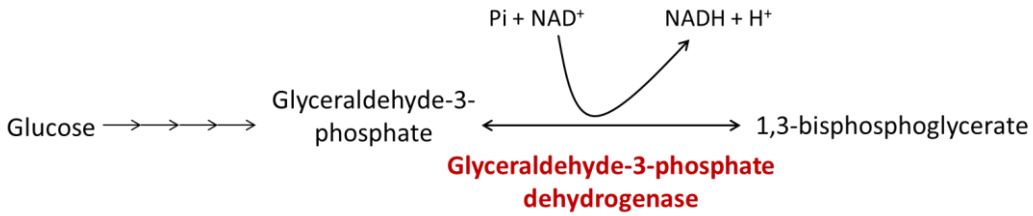

Figure 1.4. Glycolysis metabolic pathway in which Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyzes the conversion of glyceraldehyde3-phosphate (G3P) into 1,3-bisphosphoglycerate (BPG).

This step is at the center of glycolysis within cancer cells and is the convergence point of feeder and alternative ATP-generating pathways from carbohydrates. This contrasts with normal differentiated cells, where glycolysis is predominately regulated by other upstream rate-limiting enzymes. GAPDH expression is also enhanced in several cancer types and can be used as prognostic marker, since it is correlated with increased aggressiveness, drug resistance and poor progonisis.[29] Consequently, this enzyme is a putative therapeutic target, despite its widespread distribution *in vivo*.[30]

There have been a few GAPDH inhibitors reported so far (Figure 1.5), with koningic acid (KA) being one of the most studied.[8] KA, which binds covalently to the cysteine in the GAPDH active site, has been demonstrated to inhibit cancer cell proliferation.[31,32] Off-target reactivity with DNA and DNA polymerase, on the other hand, has limited its use in cancer therapy.[33] The alkylating chemical 3-bromopyruvate is another well-studied GAPDH inhibitor. Although capable of inhibiting GAPDH, the lack of selectivity for the thiols within GAPDH implies that side reactivity with other important enzyme thiols is common, limiting its utility.[34] DC-5163, 3,4-dihydroxyphenylacetaldehyde (DOPAL), and 2-phenoxynaphthalene-1,4-dione are some other GAPDH inhibitors.[34] However, insufficient inhibition combined with strong side reactions has limited the applicability of these compounds, requiring the development of novel GAPDH inhibitors.

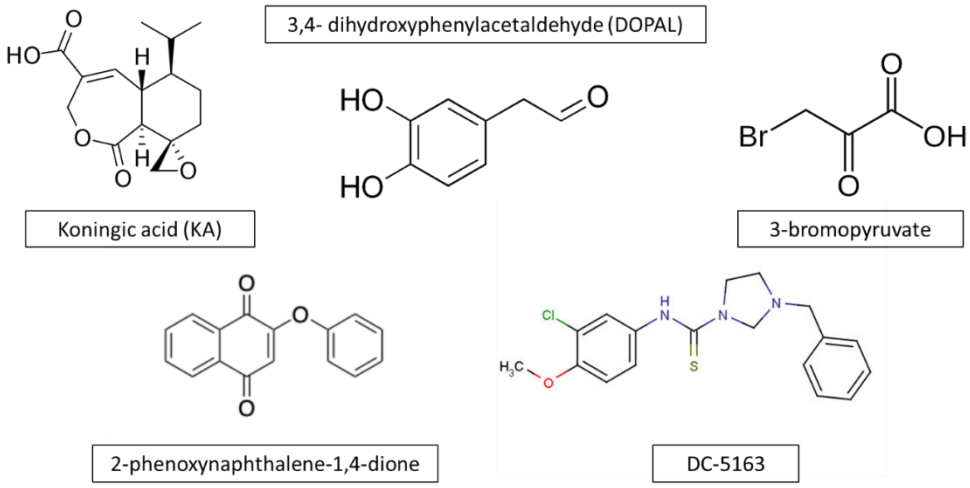

Figure 1.5. Chemical structures of previously reported GAPDH inhibitors.

In response to this, the selective modification of GAPDH by monosubstituted cyclopropenones (CPO) was investigated further in more detail. Future directions for specific GAPDH molecular targeting in cancer will depend critically on the availability of an efficient yet selective GAPDH inhibitor. In this context, it is important to establish minimum target sites that are sensitive but crucial for the enzymatic function to design and produce an effective and specific small molecule inhibitor of GAPDH. Knowledge of the high resolution three-dimensional structure of human GAPDH with specific inhibitors would facilitate the development of new anti-apoptosis drugs and, hopefully, may also represent a major advance in future cancer therapy.[27,35]

1.4. Site-Selective Protein Modification

Proteins function as one of nature's most valuable biomolecules.[36] Understanding the mechanisms of action that regulate their structure–function relationship is critical for key pharmacological targets.[37] Complex tertiary structures, which are designed for specialized purposes, make proteins great candidates for studying natural systems and developing new biological tools.[38] Proteins control a variety of basic biological activities with great specificity, both intra- and extracellularly, thanks to the combination of a highly regulated primary sequence and a complex tertiary structure.[39] As a result, proteins have been a popular target for scientists interested in studying and controlling biological processes. Due to difficulties in both three-dimensional structure prediction, labeled the "protein folding problem," and synthetic protein folding, early efforts in *de novo* protein synthesis were mostly unsuccessful.[40] Regardless of the fact that computational technologies and massive protein structure databases have significantly

improved protein structure predictions, the use of probes and chemical techniques such as modification of natural or unnatural (UAAs) amino acids[41], to modify pre-existing proteins found in nature have shown to be the most promising so far.[38,42] This approach is referred to as protein bioconjugation.

Protein bioconjugation, or the process of coupling a protein to another molecule via a covalent bond to generate a modified protein, is an emergent field of study with potential applications ranging from basic scientific discoveries to diagnostics and treatments. To make these applications a reality, the conjugation must first take place without compromising the protein function. To avoid protein denaturation, changes must take place under moderate aqueous conditions, at physiological pH, and at room temperature.[43] Furthermore, given the sensitive connection between a protein's tertiary structure and its capacity to function, modifications should lead in no alteration in the protein's tertiary structure, preserving its functional integrity.[44] Finally, bioconjugations should generate homogenous products to minimize ambiguity and increase the chances of success in their domains of application. This is particularly relevant in medical applications, where a combination of species with variable drug loading leads to changing pharmacokinetic characteristics and, as a result, a limited therapeutic window.[45]

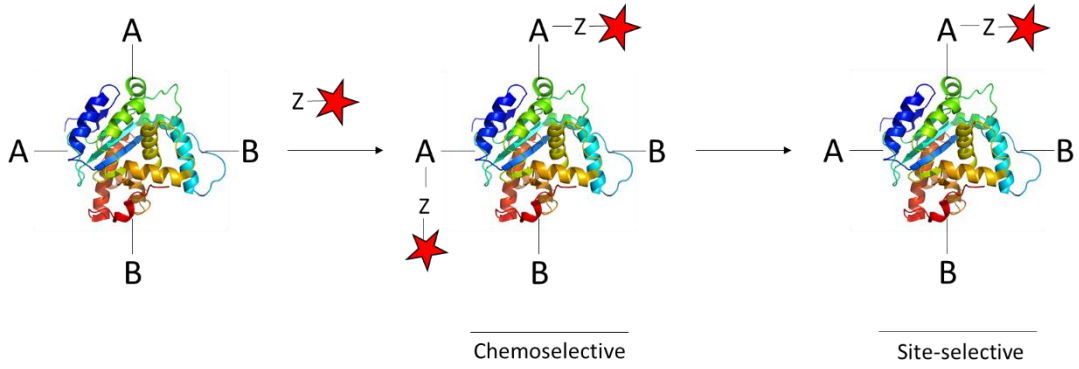

Figure 1.6. Schematic representation of the difference between chemoselective and site-selective protein modifications.

Given that the size and number of alterations have the greatest impact on product homogeneity and functional integrity, it is preferable to perform a single, minor protein modification. Considering there are often multiple instances of the same amino acid in a protein sequence, chemoselectivity alone is frequently insufficient. Therefore, during the last two decades, researchers have focused on developing chemoselective and regioselective techniques in a field known as site-selective protein modification (Figure 1.6).[44] The strategies employed to achieve site-selective modification can be split into two major categories: direct modification of native proteins and protein modification via genetic manipulation.[46] Genetic engineering is now the technique of choice for incorporating a new amino acid (canonical or noncanonical) within the structure of a protein for subsequent functionalization, and it is one of the most successful strategies for generating homogenous products. Direct native protein modification, on the other

hand, appears to be a potential developing technique, with many more cases being documented.[25,32,33]

1.4.1. Cyclopropenone-based probes

The ability to investigate the functions of specific proteins in cellular programmability and development fate through site-specific alterations of proteins or peptides is a powerful tool.[48] Targeting highly nucleophilic residues on protein surfaces is a common method of protein modification. The most commonly used functional groups for bioconjugation are nucleophilic functional groups such as the N-terminal α-amino group of peptides and proteins, the thiol group of cysteine, and the ε-amino group of lysine. [48]

Chemical probes are useful tools for studying N-terminal alterations, protein-protein interactions, protein localization, and protein dynamics, as well as other structural and functional aspects. Because of their strong nucleophilicity, cysteine and lysine residues are the most commonly modified. However, because these residues might appear numerous times in a protein at the same time, regioselective labelling can be challenging. [49] N-terminal cysteine appears to be a one-of-a-kind practical solution for overcoming regioselectivity and site-specificity issues. [50,51]

Due to their small size and unique mode of reactivity, cyclopropenones (CPO), are considered appealing motifs for bioorthogonal chemistry. However, their application has remained largely unexplored in the field. Since its discovery in 1959, CPO has been used in a wide range of synthetic organic applications.[52] Comprising a highly strained 3-membered ring and a strong dipole moment induced by a carbonyl group, monosubstituted cyclopropenones have recently been shown to react selectively with N-terminal cysteine residues in proteins containing multiple cysteine residues under mild, biocompatible conditions.[49]

However, the precise understanding of the CPO reactivity is yet to be fully explored and there is potential for the reaction scope and utility of CPO to be expanded. If successful, the novel CPO-specific motif could be genetically incorporated into proteins and therefore provide a generic, protecting-group-free strategy for protein modification.

1.5. Three-dimensional structures of GAPDH

In the past two decades, GAPDH, a glycolytic enzyme that was long thought to be a simple “housekeeping” protein, has been shown to be engaged in a variety of cellular activities other than glycolysis. GAPDH is typically found in the cytoplasm as a tetrameric isoform made up of four identical ~ 37 kDa subunits, each with a single catalytic thiol group.[18]

Since there is a rising interest in GAPDH as a target for developing chemotherapeutics, the determination of the three-dimensional structure of GAPDH was an important contribution to

structure-based drug design. The first human GAPDH crystal structure was reported by Watson and colleagues at 3.5 Å in 1976, determined from twinned crystals of the human skeletal muscle enzyme.[53] Since then, many other structures of human GAPDH have been reported, namely by Ismail & Park at 2.5 Å in 2005 (PDB code 1znq)[35] or by Jenkins and Tanner at 1.75 Å in 2006 (PDB code 1u8f)[54], which is the highest resolution to date of a GAPDH crystallographic structure.

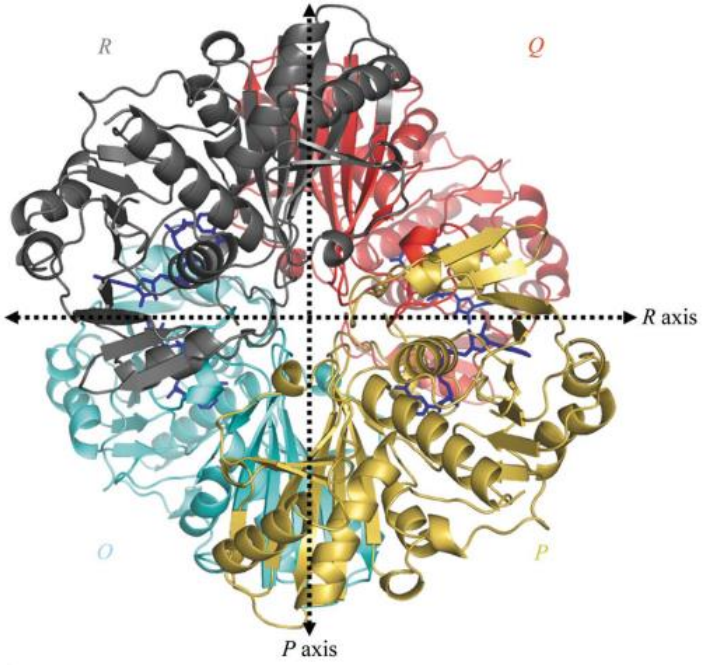

Figure 1.7. An overview of the human liver GAPDH homotetramer. Subunits O, P, Q, and R are depicted in cyan, yellow, red, and gray, in that order. Each monomer is attached to a NAD+ molecule (blue). Symmetry axes are depicted as dashed arrows, and the Q axis is perpendicular to the paper's plane. (Adapted from [35]; PDB code 1znq).

The overall structure of the homotetrameric human GAPDH consists of four subunits designated O, P, Q and R where each subunit is bound to a NAD+ molecule. Each subunit consists of an N-terminal NAD+-binding domain (residues 1-150) and a C-terminal catalytic domain (residues 149-313). The N-terminal NAD+-binding domain has an α/β dinucleotide-binding fold, while the C-terminal catalytic domain consists of an eight-stranded mixed parallel β-sheet, where the strands are connected by either short α-helices or turns (Figure 1.7).[35] The active site of each subunit is located in a large cleft between the NAD+ -binding and catalytic domains. The Cys nucleophile (Cys152) resides at the N-terminus of the first helix in the catalytic domain.[35]

The creation of GAPDH antagonists for the treatment of cancer or cancer-related disorders has recently become the focus of many structure-based drug design investigations. Over the years, the search for new GAPDH antagonists has been growing, and as a consequence, a crystallographic structure of these antagonists in complex with the protein has become extremely important.

So far, there are two structures of GAPDH in complex with inhibitors that are worth mentioning (see Figure 1.8). The first one was reported by Guido *et al*. in 2009, and demonstrates the binding mode of iodoacetate (IAA), an alkylating agent of several amino acid residues, to the active site of GAPDH. It inactivates a variety of enzymes by acting on their catalytic cysteine residues. Accordingly, the crystallographic data shows that the cofactor molecule ($NAD^{+)}$ plays a central role in the binding of the inhibitor within the GAPDH active site (PDB code 3DMT)[55]. The other study was reported by Park *et al*. in 2019 (PDB code 6IQ6)[56], in which they investigated the differences between IAA and di- or monomethyl fumarate (DMF or MMF) inhibitord. DMF is a well-known oral immune-modulator used to treat relapsing-remitting multiple sclerosis (RRMS) and psoriasis. It works by limiting hGAPDH-related aerobic glycolysis and inhibiting the catalytic activity of hGAPDH. Furthermore, DMF covalently modifies the catalytic cysteine (Cys152) of hGAPDH via succination, irreversibly inactivating the enzyme. While the inhibitory mode of IAA allowed the binding of NAD^+, MMF, due to its bulkier size (compared to IAA), inhibited the access of NAD^+.[56]

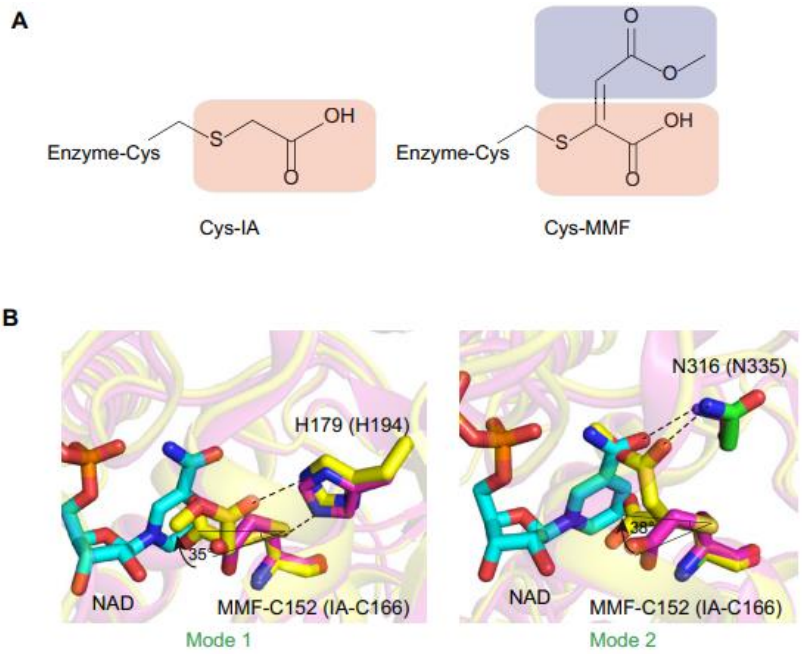

Figure 1.8. Differences of binding modes to GAPDH between MMF and iodoacetate. (A) Iodoacetate and MMF chemical structures. The red and purple round-squares represent the inhibitors' similar and distinct chemical structures, respectively. (B) The difference in inhibitory modes between iodoacetate and MMF. (Adapted from [56]).

The crystallographic structures are extremely important for the study of the detailed inhibitory mode of different antagonists of GAPDH and it's necessary to explain the differences between inhibitory efficacies. These structural data provide insights for the design and development of new structure-based drug design that can be used for the treatment of cancer or cancer-related disorders.

1.6. Production of recombinant hGAPDH

1.6.1. Expression of hGAPDH in *E. coli* and purification using a Strep Tag

Over the years, the relevance of the study of GAPDH in cancer therapy and in the development of new drugs has led to more and more new production and purification protocols for this protein.

Escherichia coli (*E. coli*) is commonly used to express a wide range of heterologous proteins. Efforts have been undertaken to increase the desired protein expression level. However, difficulties with protein expression regulation persist, necessitating the development of novel ways to address these concerns.[57] Despite the many possible hosts for the production of human GAPDH, the most used and the most efficient is the production in *E. coli*.[35,54]

In contemporary protein research, efficient purification, detection, and immobilization or separation procedures—possibly in complex with cognate macromolecules—are critical. Strep-tag II is an eight-residue peptide sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) with inherent streptavidin affinity that may be fused to recombinant proteins in a variety of ways.[58] As a result, the Strep-tag II peptide, which is short, physiologically inert, proteolytically stable, and does not interfere with membrane translocation or protein folding, is a useful tool for both quick separation of a functional gene product and detection or molecular interaction analysis.[58]

1.7. Scope of the thesis

The main goal of this thesis is to obtain a functional and structural characterization of GAPDH in complex with a CPO-based probe. The growing relevance in the impact of GAPDH in cancer cells has encouraged the design of more potent and selective inhibitors with improved pharmacological properties to overcome the already existed ones. Thus, the Gonçalo Bernardes group, located at the University of Cambridge, has developed an irreversible GAPDH inhibitor, utilizing cyclopropenones (CPOs) as selective inhibitors and as components for performing bioorthogonal reactions on proteins. In an exciting development, preliminary unpublished work by the Bernardes group demonstrated that monosubstituted CPOs can selectively modify N-terminal cysteines on a protein that also contains internal cysteine residues.[59] Since the molecular basis underlying CPO-based probe inhibition is still not fully understood, further elucidation of the functional and structural characteristics of GAPDH alone or in complex with these irreversible inhibitors are necessary for feasibility testing of therapeutic efficacy. X-ray crystallography can also provide additional information regarding a molecular view of the GAPDH-NAD^+ binding pocket and how it binds to different inhibitors.

To fulfill these requirements, the work of this master thesis aimed at:

- producing recombinant human GAPDH in *E. coli*,
- establishing a protocol for protein purification using a Strep-tag,
- undertaking a biochemical and biophysical characterization of the purified protein, alone or in complex with the CPO-probe,
- Undertaking in parallel crystallization trials to determine the 3D structure of the protein alone and in complex with the CPO-probe.

In order to obtain sufficient material to perform crystallization trials it's of most importance to develop an expression vector with the structural gene for human GAPDH (Uniprot p04406). The synthetized construct will be used for bacterial expression using a N-terminal Strep-tag to facilitate protein purification. Next, different biochemical and physical approaches such as UV-spectroscopy, polyacrylamide-based gel electrophoresis (SDS and Native-PAGE), Western Blot, Dynamic Light Scattering, Circular Dichroism and LC-MS will be carried out to assess the quality of the protein sample for crystallization. Differential Scanning Fluorimetry will be used to characterize GAPDH in complex with the CPO-probe. In a final step, purified hGAPDH will be co-crystallized with the CPO-probe inhibitor and crystallization experiments will be optimized in an attempt to obtain high-quality crystals for X-ray diffraction.

This Master plan intends to explore CPO-probes as selective inhibitors of GAPDH and elucidate the binding mode of these inhbitors to the target protein. Hopefully, this may also represent a major advance in future cancer therapy.

Chapter 2

2. Methodology

2.1. Genetic constructions

The synthetic human GAPDH gene construct was designed and ordered from TwistBiosciences™. The insert included a ribosome binding site and ATG start codon, an N-terminal StrepTag II, a Tobacco Etch Virus (TEV) protease cleavage site fused to the gene sequence of human GAPDH (StrepTag II-TEV-hGAPDH). The plasmid was a pET-21(+) also contained one antibiotic resistance gene (ampicillin), two cloning sites of EcoRI and NotI restriction enzymes, and a T7 promoter for high-level gene expression in bacterial strains (see Figure 2.1).

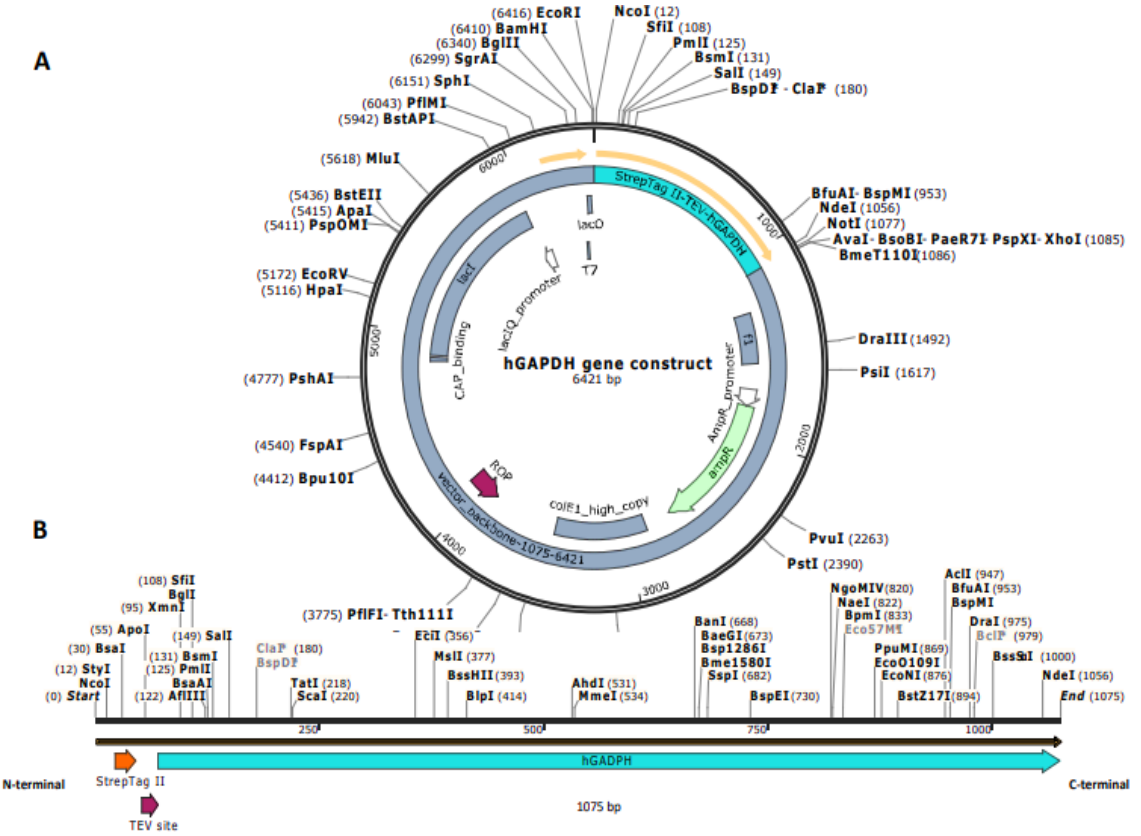

Figure 2.1. Recombinant plasmid map from pET21(+) (**A**) containing the StrepTag II-TEV-hGAPDH gene construct sequence (**B**).

For DNA cloning or plasmid amplification purposes, NZ5alpha cells (NZYTech) were routinely used. Cultures were grown in Luria's broth (Miller's formulation) low salt medium (Sigma, Cat# L3397) with (solid) or without (liquid) 10 g/L agar. All plasmids used in this work coded for an Ampicillin resistance cassette. Either ampicillin (solid cultures) or carbenicillin (liquid cultures) were incorporated in the culture medium at 100 µg/mL and used as selective agents. Cultures were routinely incubated at 37 ºC, in a standard bacteria oven (solid cultures), or different orbital shakers, according to the experiment.

2.2. Transformation of competent *E. coli* cells

Plasmids were used to transform chemically competent cells using the heat-shock method. Briefly, DNA was incubated with cells (<10% v/v of DNA solution in DNase/RNase-free MilliQ water). Tubes were swirled gently and incubated on ice for 30 min. After incubation, cells were heat-shocked at 45 °C in a block heater for 45 sec and then immediately placed on ice for 2 min. Then, Luria-Bertani medium (LB) was added to the cells and incubated at 37 °C for 30 min with shaking at 225-250 rpm. Tubes were pelleted and the resuspended pellet was plated on LB agar plates containing ampicillin and incubated at 37 °C.

2.3. Cell growth

A single colony of previously transformed cells was pre-inoculated into LB-medium supplemented with carbenicillin. The culture was left growing at 37 °C to an optical density (OD_{600}) of 0.6. Subsequently, fresh LB-medium supplemented with the appropriate antibiotic was inoculated with 1.5 % of pre inoculum. The culture was left growing at 37° for approximately 4 h, to a final OD_{600} of 0.8.

2.4. Expression and protein purification

For protein expression purposes, BL21(DE3) cells were routinely used. After the hGAPDH gene construct was introduced into the expression strain BL21(DE3) using the same method as described in the previous section, the protein was expressed by induction with 0.4 mM of isopropyl β-d-thiogalactopyranoside (IPTG), and the culture was left to incubate with shaking at 37 ˚C till it achieved an optical density of 0.8. Induction with IPTG lasted for 16 h at 18 °C, 120 rpm.

Collection of bacterial cells was performed on 250 mL centrifuge-appropriate flasks, at 4 °C, on a J2-21M/E High-Speed Centrifuge, equipped with a JA-14 rotor, at 8000 x g, for 20 min.

Pelleted *E. coli* cells were resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5; 300 mM NaCl; 1 mM EDTA, 1 mM DTT; DNase and Protease inhibitor (Roche, cOmplete™, Mini Protease Inhibitor Cocktail)). The resuspended cells were disrupted by sonication, on a Soniprep 150, in an ice-chilled container for 4 min (20'' ON + 20'' OFF). After centrifugation for 15 min, at 18 000 x g and 4 ºC (Avanti J-25 High Speed, Rotor JA 25-50) we obtained the soluble fraction (SF) and a pellet. Then, the pellet that consisted of inclusion bodies (IB) was again washed with lysis buffer adding 1% Triton X-100. The resuspended pellet was sonicated for 4 min (20'' ON + 20'' OFF) and was centrifuged for 25 min, at 18 000 x g, at 4 ºC. After centrifugation, the supernatant was collected (Washes) and the pellet containing the IB again was resuspended again in lysis buffer + 0.5% Triton X-100, mixed by vortexing for 1 min and centrifuged as described above. Finally, we discarded the pellet (IB), collected the second supernatant, and added it to the previously

isolated Washes. At the end, two fractions were obtained: a soluble faction (SF) and the IB washes (Washes), that were filtered via a 0.45 µm syringe filter and kept on ice for further purification.

The purification process was carried out first with the SF. The SF fraction was loaded into an affinity-flow column for Strep-Tactin®, following the protocol provided.[60] Four column volumes of wash buffer (150 mM NaCl, 100 mM Tris-HCl, 1 mM EDTA, pH=8) were applied to the packed column and the protein was eluted with 150 mM NaCl, 100 mM Tris-HCl, 1 mM EDTA, pH 8 with 2.5 mM desthiobiotin. The same procedure was applied to the Washes fraction. All fractions from both purification processes were analyzed by SDS–PAGE gel.

After purification, all fractions eluted were loaded on a PD-10 desalting column (GE Healthcare). These prepacked Desalting columns are designed for rapid, convenient sample clean-up of proteins and other large biomolecules (>5000 M_w). In this particular case, the PD-10 desalting column was used to exchange the protein buffer from Tris-HCl pH 7.5 and 30 mM NaCl to 20 mM Phosphate buffer pH 7.4 and 30 mM NaCl. Buffer exchange of hGAPDH was performed following the GE Healthcare protocol.[61]

2.5. SDS-polyacrylamide gel electrophoresis

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out in order to separate the proteins according to their electrophoretic mobility.[62] Concentration of the resolving gel (between 10 % and 15 %) was chosen according to the size of the protein. The resolving gel contained the separating buffer, composed of 1.5 M Tris pH 8.8 and 30% acrylamide/bisacrylamide solution, (Invitrogen, Karlsruhe) whose volume depended on concentration of the gel, 10% SDS (w/v), ddH_2O, 10 % APS (ammonium persulfate, Merck, Dresden) and 5 µL TEMED (Sigma-Aldrich, Deisenhofen). 5 mL of the resolving gel mixture was poured between gel plates (10.5 cm x 9.8 cm, 1 mm spacer) and was polymerized in the presence of isopropanol. After 30 minutes isopropanol was removed, 2 ml of stacking gel mixture at 4% (30 % acrylamide/bisacrylamide solution, 0.5 M Tris, 10 % SDS (w/v), ddH_2O, 10 % APS and 5 µL TEMED, was poured, silicon combs inserted, and gel left to polymerize. Depending on the protein concentrations, 5 µl to 15 µl of the samples were loaded per lane. In addition, 5 µl of the NZYBlue Protein Marker (NZYtech) loading buffer were loaded for size estimation. Electrophoresis was performed in SDS running buffer at 100 mA for about 45 min. Proteins were visualized by Coomassie staining with InstantBlue™ (Expedeon).

2.6. Western blot analysis

Following separation of proteins by SDS-PAGE, proteins were transferred onto a PVDF membrane as described next. Transfer cassettes contained the following layers: a sponge,

Whatman 3MM blotting paper, PVDF membrane (Pall corporation), SDS-PAGE gel, Whatman blotting paper and a sponge (all equipment was pre-immersed in methanol-free transfer buffer (25 mM Tris base, 190 mM glycine, pH 8.3). The transfer cassette was then placed in a transfer tank filled with transfer buffer for 1 hour at 350mA. Following the transfer, PVDF membranes were blocked with 5% milk powder in Tris-buffered saline with 0.1% Tween20 for 1h and immunodetection was performed with an antibody for StrepTag II, following to the manufacturer protocol.

2.7. Dynamic light scattering

Dynamic light scattering (DLS), is one of the most important experimental techniques currently available to determine the size distribution, stability, and aggregation state of small particles in suspension.[63] Detection of light scattering from matter is a useful technology, with countless applications in research. According to the light source and detector, specific properties of molecules can be studied.

DLS has become a routine method in crystallography, mostly used to evaluate the possibility of protein crystallizability due to the fact that is easy to implement, is quick to perform, and is least destructive to the sample.[64]

DLS was performed to obtain a characterization of particle size measurements and to detect aggregates in purified samples. This experiment was performed on a SpectroLight™ 600 (XtalLight ™200/210, Xtal concepts). Two 20 µL samples were prepared, at concentrations of 1.5 mg/mL and 0.75 mg/mL. 1 µL of each solution was added to a 96-well plate. Measurements were set to run for 2 hours and were carried out at 4°C. Data collected by the DLS software was provided in terms of hydrodynamic radius distribution along time along with a statistical analysis.

2.8. Differential scanning fluorimetry

Differential Scanning Fluorimetry (DSF), also known as ThermoFluor or Fluorescence Thermal Shift Assay (FTSA) is a simple, fast, inexpensive, and widely applicable biophysical method that is typically used to study the thermal stability of proteins under various conditions.[65]

In DSF, protein stability is determined by gradually increasing the temperature and measuring the activation of a fluorescent dye that binds to the hydrophobic parts of the protein as it unfolds.[66,67]

2.8.1. RUBIC Buffer Screen

DSF Rubic Buffer Screen experiments were carried out to measure the change in protein stability in a range of temperatures. The assay was performed in MicroAmp™ EnduraPlate™ Optical 96-Well Clear Reaction Plates with Barcode (Applied Biosystems, Life Technologies, California, USA) using a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). The final reaction mixture (20 µL of total volume) contained 4 µg of GAPDH, 4-fold of Protein Thermal Shift™ Dye (Applied Biosystems) diluted in protein buffer solution. The temperature was increased from 25ºC to 90ºC with an increment rate of 0.016ºC/s. Excitation and emission filters were applied for Protein Thermal Shift™ Dye (470nm and 520nm, respectively) and for ROX reference dye (580nm and 623nm, respectively). The melting temperatures were obtained by taking the midpoint of each transition.

2.8.2. Thermo Shift Assay with GAPDH-CPO complex

DSF experiments were carried out to measure the change in protein stability upon binding with inhibitor compounds. The assay was performed in MicroAmp™ EnduraPlate™ Optical 96-Well Clear Reaction Plates with Barcode (Applied Biosystems, Life Technologies, California, USA) using a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). Pre-incubation of the protein with CPO-probe for 2 hours at 4ºC was required prior to DSF experiments. The final reaction mixture (20 µL of total volume) contained 4 µg of hGAPDH, 4-fold of Protein Thermal Shift™ Dye (Applied Biosystems) diluted in protein buffer solution, and three different concentrations of the CPO-probe; 0.250 mM, 0.5 mM and 2 mM. The temperature was increased from 25ºC to 90ºC with an increment rate of 0.016ºC/s. Excitation and emission filters were applied for Protein Thermal Shift™ Dye (470nm and 520nm, respectively) and for ROX reference dye (580nm and 623nm, respectively). The melting temperatures were obtained by taking the midpoint of each transition.

2.9. Circular Dichroism Spectropolarimetry

Circular dichroism (CD) spectropolarimetry is an optical spectroscopic method that refers to the differential absorption between the two circular components of plane-polarized light – one rotating counter-clockwise (left-handed, L) and the other clockwise (right-handed, R) – by optically active molecules, in order to determine their absolute configurations.[68,69]

The CD spectra of the protein were acquired using a J-815 CD spectrometer (Jasco, Easton, USA) in a 200-250 nm wavelength range. Protein buffer exchange to PBS was carried out prior to the CD experiment, using a PD-10 column. Samples of Apo-GAPDH and GAPDH in conjugation with CPO-probe (in ACN and DMSO at 2-fold) were diluted to 0.3 mg/mL, and 200 µL were transferred to absorption cells of 1 mm path length (Hellma Analytics, Mülheim, Germany). Experiments were made at 25 ºC and the final spectrum was an average of three

consecutive scans, subtracted by a blank spectrum. Collected data was normalized by adjusting the values of ellipticity (θ) to molar concentrations.

2.10.GAPDH activity assay

The activity of pure GAPDH protein was assessed according to the instruction of the GAPDH activity assay kit.[70] To obtain a NADH Standard for colorimetric detection, we start by adding 0, 2, 4, 6, 8 and 10 µL of 1.25 mM NADH Standard into a 96 well plate to generate 0 (blank), 2.5, 5.0, 7.5, 10 and 12.5 nmol/well of NADH Standard. Well volume was set to 50 µL with GAPDH Assay Buffer. The reagent background for the assay is the value obtained for the 0 (assay blank) NADH Standard. A new standard curve must be set up each time the assay is carried out.

Two GAPDH activity assays were performed. The first one in which two different concentrations of CPO-probe were tested, 35 µM and 350 µM, using three different incubation times: 15, 30, and 60 min. The second assay (control) was intended to verify the effect of DMSO on GAPDH activity, using probe concentration of 35 µM and incubation times of 30 and 60 min. For both assays, we added 50 µL of hGAPDH to each well and then added the probe at the desired concentration and respective incubation times. For the control, DMSO was added instead of the probe, at the same concentration (v/v) and same incubation times. Finally, we added 50 µL Reaction Mix (which contained 46 µL GAPDH Assay Buffer, 2 µL GAPDH Developer and 2 µL GAPDH Substrate) to each sample and standard control well.

Final mixture was incubated at 37 °C for 60 min, absorbance of each sample at 450 nm was measured at 0 min and 60 min, and thence, the GAPDH activity was calculated according to the absorbance values and NADH Standard Curve.

2.11. Sturcture determination by X-ray crystallography

2.11.1. hGAPDH crystallization and hGAPDH co-crystallization with CPO-probe

High-throughput screenings were set up to test two of commercial crystallization screens, the Structure 1 and 2 and BCS (both from Molecular dimensions Ltd., Suffolk, UK). The trials were carried out using the sitting-drop vapor diffusion method with the mosquito® LCP crystallization robot (TP Labtech Ltd, Hertfordshire, UK). The purified hGAPDH was used for crystallization at a final concentration of 10 mg/mL (according to ref. [71]). The drops consisted of 0.100 µL of the reservoir solution mixed with an equal volume of the protein sample and equilibrated against a 45 µL reservoir. Plates were incubated at 20ºC. Promising conditions that favored crystal formation were used for scale-up and the best-shaped crystals were analyzed at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France and at ALBA Synchrotron in Barcelona, Spain.

The same protocol described above was repeated with the hGAPDH-CPO complex. The purified hGAPDH was used for crystallization at a final concentration of 10 mg/mL (according to ref. [71]) and pre-incubated for 2 hours at room temperature with a 2-fold concentration of CPO-probe in DMSO.

2.11.2. Model building and refinement

An X-ray diffraction data set to 2.11 Å was collected at ESRF beamline ID30A-3 with a Dectris EIGER X 4M detector from a cryocooled crystal at 100 K. The diffraction data were processed with AutoPROC[72] and XDS[72], and the data processing and refinement statistics are summarized in Table 2.1. Two diffraction datasets were obtained: in the first, a spherical region of reciprocal space to 2.2 Å resolution was defined, and in the second a triaxial ellipsoidal region to a maximal resolution of 1.95 Å was selected with the STARANISO module of AutoPROC[72]. The structure of hGAPDH was determined by molecular replacement with PHASER[73] as implemented in the CCP4 program suite[74,75] using the PDB entry 1znq[35] as a search model, without including ligands and water molecules. Three independent copies of the search model were located in the crystal structure, and model rebuilding was carried out with BUCANEER[76] and COOT[77]. Initial structure refinement was undertaken with REFMAC[78].

Refinement was continued with PHENIX[79], alternating with manual model editing in COOT between refinements against σA-weighted $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps. Hydrogen atoms were added in calculated positions, water molecules were added automatically, and the final refinement was carried out to 2.11 Å against the STARANISO dataset, where positional and isotropic thermal motion parameters of the non-hydrogen model atoms were refined. The statistics are included in Table 2.2. Figures were prepared with PYMOL[80].

Table 2.1. Data processing statistics for the crystallographic structure of hGAPDH.

Data Collection	**hGAPDH**
Beamline	ESRF ID30A-3
Detector	Eiger X 4M
Wavelength (Á)	0.96770
Space Group	I 222
Unit cell parameters:	
a, b, c (Á)	90.89 130.25 220.71
Data Processing	AutoPROC / STARANISO
Resolution limits of ellipsoid fitted to resolution cut-off surface (Å)	2.36, 2.06, 2.78
Resolution, spherical limits (Å)	74.5– 2.11 (2.33-2.12)
Nr. Observations	264508 (13189)
Unique reflections	49416 (2472)
Multiplicity	5.4 (5.3)
Completeness, spherical (%)	66.3 (13.4)
Completeness, ellipsoidal (%)	93.5 (66.1)
R-merge (%) [b]	9.4 (110)
R-p.i.m. (%) [c]	4.5 (51.3)
<I/σ (I)>	13.1 (1.6)
CC $^{1/2}$	0.998 (0.607)
Wilson B (Å^2)	54.3
Z [d]	3
Estimated V_M [e]	3.00
Estimated Solvent Content (%) [e]	59.0

[a] Values in parentheses refer to the highest resolution shell; [b] R-merge = merging R-factor, (Σ_{hkl} Σ_i $|I_i(hkl) - \langle I(hkl)\rangle|$) / ($\Sigma_{hkl}$ Σ_i $I(hkl)$) $\times$ 100 %; [c] R-p.i.m. = precision-independent R-factor, Σ_{hkl} $[1/(N-1)]^{1/2}$ Σ_i $|I_i(hkl) - \langle I(hkl)\rangle|$ / (Σ_{hkl} Σ_i $I_i(hkl)$) $\times$ 100 %.[81] For each unique Bragg reflection with indices (hkl), I_i is the i-th observation of its intensity and N its multiplicity; [d] Nr. molecules in the asymmetric unit; [e] According to [82].

Table 2.2. Refinement statistics for the crystallographic structure of hGAPDH.

Dataset	hGAPDH
Resolution limits (Å)[a]	74.54 – 2.12 (2.16 – 2.12)
% R_{work} [b]	16.4 (29.7)
% R_{free} [c]	21.2 (34.4)
ML coordinate error estimate (Å) [d]	0.23
Model composition and completeness	
Regions omitted [e]	1A-5A
Non-hydrogen protein atoms [f]	7688
Ligand/ion atoms	132
Solvent molecules	270
Mean B values (Å^2) [g]	
Protein	47.0
Ligand/ion	47.8
Solvent	43.1
Model r.m.s. deviations from ideality	
Bond lengths (Å)	0.012
Bond angles (°)	1.216
Chiral centers (Å^3)	0.063
Planar groups (Å)	0.009
Model validation [h]	
% Ramachandran outliers	0.3
% Ramachandran favored	95.9
% Rotamer outliers	2.08
% C^{β} outliers	0
Clash score	4.05

[a] Values in parentheses refer to the highest resolution shell; [b] $R_{work} = (\Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)||) / (\Sigma_{hkl} |F_{obs}(hkl)|) \times 100$ %; [c] R_{free} is calculated as above from a random sample containing 5% of the total number of independent reflections measured; [d] Maximum-likelihood estimate by PHENIX[83]; [e] Chains B, D, F and H correspond to the large subunit HysA; chains A, C, E and G correspond to the small subunit HysB; [f] Including atoms in the alternate conformations of disordered groups of residues; [g] Calculated from isotropic or equivalent isotropic B-values; [h] Calculated with MolProbity[84].

2.12.Cell biology

2.12.1. Cell lines

MCF7 cells and MDA-MB-231 cells (human breast cancer cells), A549 cells (human lung cancer cells), HepG2 cells (human liver cancer cells) and HEK-293 cells (human embryonic kidney cells) were cultured in DMEM containing 4.5 g/L glucose, 10% FBS, HEPES, Non-essential amino acids (NEAA), 5% CO, Glutamax and penicillin/streptomycin (Sigma). All cells were cultured at 37 °C.

2.12.2. Cell proliferation and IC_{50} determination

One hundred microliters of culture medium containing 10 x 10^3 of each cell line, MCF7 cells, A549 cells, MDA-MB-231 cells, HEPG2 cells or HEK cells, was added to 96-well plate wells. The plates were incubated at 37 °C for 24 h for seeding, and then cells were treated with different concentrations of CPO-probe (DMSO 100% was used as a control). Triplicates were made. After treating for 48 h, the medium was discarded and replaced with fresh medium containing CellTiter-Blue® (CellTiter-Blue® Cell Viability Assay, Promega) (10 µL of CellTiter-Blue® per 190 µL of medium). After incubation for 1 – 2 h at 37 °C, the fluorescence value was recorded at 560 Excitation and 590 Emission (560Ex/590Em).

2.12.3. Intracellular GAPDH activity assay

MCF7 cells, A549 cells, MDA-MB-231 cells, HEPG2 cells and HEK cells were plated in 24-well plates with 5 × 10^4 cells per well and cultured for 24 h. 5 µL of CPO-probe in DMSO at 35 µM, were then added to each well. Cells were incubated at 37 °C for 48 h, after this time they were washed three times with PBS, rapidly homogenized with 100 µl GAPDH Assay Buffer, and kept on ice for 10 min. Next, cells were centrifuged at 10,000 x g, 4 °C for 5 min and 5 µL supernatant per well for the assay was added. 0, 2, 4, 6, 8 and 10 µL of 1.25 mM NADH Standard was added into a series of wells in a 96-well plate to generate 0, 2.5, 5.0, 7.5, 10 and 12.5 nmol/well of NADH Standard. Volume was set to 50 µL/well with GAPDH Assay Buffer followed by the addition to each well of 50 µL Reaction Mix, which contains 46 µL GAPDH Assay Buffer, 2 µL GAPDH Developer and 2 µL GAPDH Substrate. The final mixture was incubated at 37 °C for 60 min. Absorbance at 450 nm was measured at 0 and 60 min from each sample, and the inhibition rates were calculated according to the absorbance values and NADH Standard Curve.[70]

Chapter 3

3. Results and discussion

3.1. Expression and purification process

The recombinant hGAPDH was expressed with a Strep-tag® II (WSHPQFEK) fused to its N-terminal for high selective binding to an engineered streptavidin ligand, called Strep-Tactin® Sepharose. This technology allowed one-step purification of hGAPDH under physiological conditions, thus preserving its biological activity. The fractions from all purification steps (Flow-through – FT – Wash 1 and Wash 2 – W1 and W2 – and Elution 1 and 2 – E1 and E2) were analyzed by denaturing SDS-PAGE gel (Figure 3.1). The fractions were also analyzed by Western Blot (Figure 3.2) and the fractions that displayed the highest purity were pooled and concentrated for subsequent experiments.

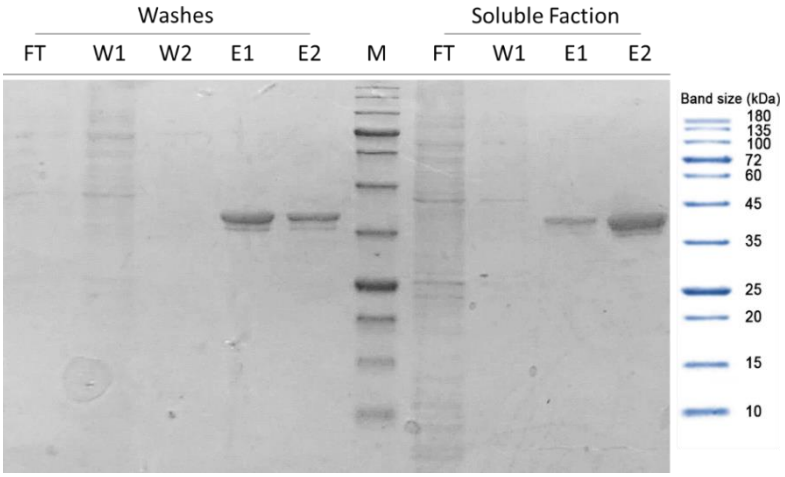

Figure 3.9. SDS-PAGE gel was run on a 15% acrylamide gel in parallel with NZTBlue Protein Marker (M).

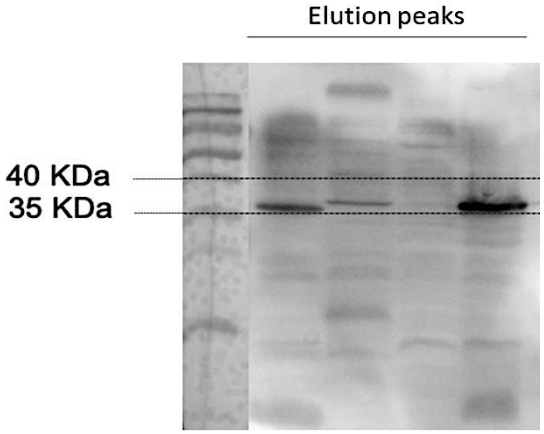

Figure 3.2. Western Blot analysis. Eluted fractions (from Washes and SF) were quantified, and 30 µg from each fraction were loaded on a 15 % sodium dodecyl sulfate−polyacrylamide (SDS) gel for electrophoresis, as indicated. The SDS gel was transferred to a PVDF and imaged.

In Figure 3.1, a single band is present in all eluted fractions, both resulting from the SF and the Washes. This band is in the 35-45 kDa range, in agreement with the molecular weight of the recombinant protein (38 KDa). After one purification step, we obtained a total of 7.72 mg of purified protein per liter of cell culture. Of these 7.72 mg of total protein, 4.0 mg were from the IB Washes eluted fractions, and 3.7 mg were from the SF eluted fractions. These results justify the extra washing step of the IB since a considerable amount of protein was recovered from this fraction. Expression of human GAPDH with a N-terminal Strep-tag II bearing a Tobacco Etch Virus (TEV) cleavage site in the linker region was successful, even using the standard *E. coli* BL21(DE3), at 16h post-induction, at 18 °C and low-agitation, in LB (Luria's formulation) medium. Longer induction times did not substantially improve soluble protein expression levels.

In all the produced batches of the fusion protein, purification from the soluble fraction using one single affinity-column was enough to obtain purified protein. Even if tag-removal would be required, the whole purification process could be performed using only one type of purification column. It has been previously described that the fusion of a N-terminal His-tag to hGAPDH complicates the purification process as more steps are needed. [54,85] In combination with high density mammalian protein expression, the His-tag system can lead to poor purification results, if conditions are not optimized. In order to achieve pure proteins for downstream applications from the coupling of the His-tag system and high density mammalian systems, further adjustments like dialysis of the supernatant becomes necessary. Such adaptions require additional optimization time and effort; however, these challenges can be avoided by using the Strep-tag® system.

3.2. Purified sample quality assessment by DLS

The reason why protein crystal formation is a major challenge is related with the large number of physicochemical and biochemical variables that impact the thermodynamics of this process. Some of the characteristics that may impact protein crystallization include homogeneity, chemical purity, stability, aggregation, and the structural state of a protein in solution. Methods like DLS and CD can offer helpful information on these factors and can aid in determining when a protein becomes more stable or begins to denature. [86]

The first DLS measurement was performed on a purified sample of hGAPDH to evaluate the dispersion level of the particles in solution and their aggregation state. The sample was previously filtered, and the experiment was carried out at 4ºC for 2 hours. The correlation function (Figure 3.3 A) displayed a typical curve for monodisperse samples and the size distribution (Figure 3.3 B) of the particles in solution was consistent with the size of the protein, meaning that the protein is stable in solution, without the presence of aggregates.

Only a population with hydro-dynamic diameter D_H= 3.99 ± 0.25 nm Å was observed, according to Figure 3.3. This observation is compatible to a D_H value for a dimer with a Mw of 78.86 KDa (Figure 3D). A D_{max} value of 100 nm was determined for the tetrameric GAPDH.[87] The

CONTIN[88] program, autocorrelation functions such as those seen in Figure 3.3 A, for example, were analyzed to obtain the distribution of particle radii, as seen in Figure 3.3 B. To study the temporal development of the reaction in the droplet requires taking a regular series of measurements over a few hours. Each series can consist of several hundred individual measurements. In Figure 3.3 C, an example is given in which the distribution of radii (horizontal axis) is seen as a function of time (vertical axis). The comparative abundance of each radius fraction is depicted in a color scale (from blue to red).

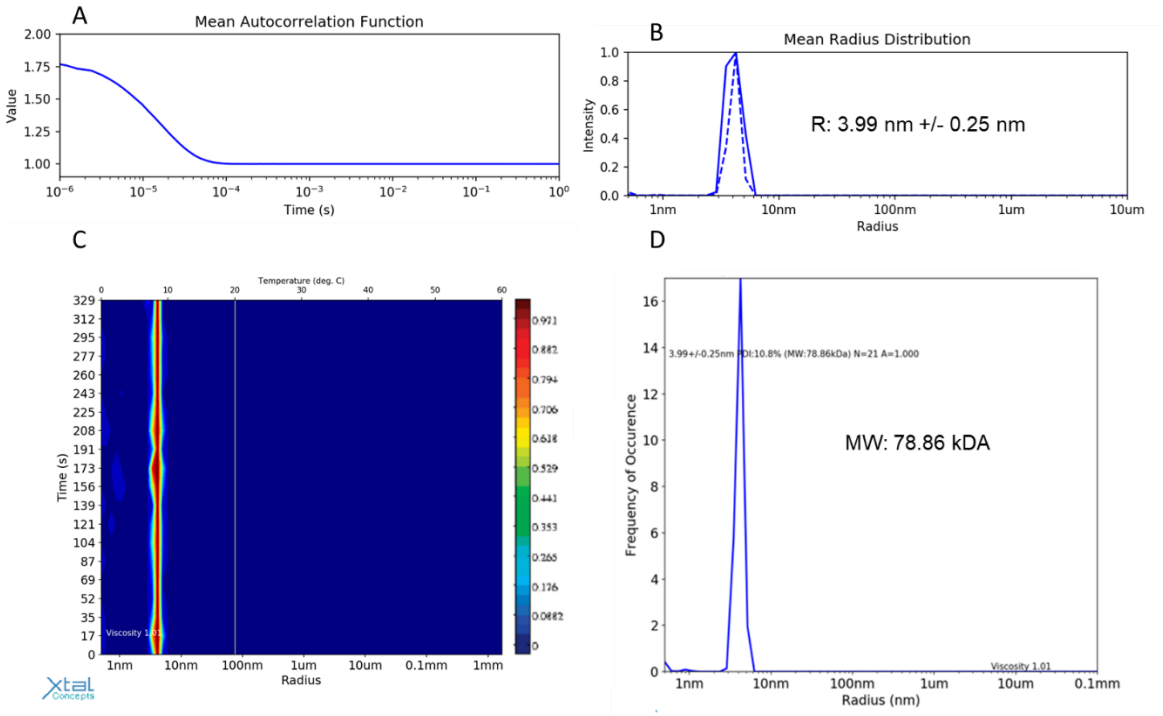

Figure 3.3. Summary hydrodynamic radius display of a measurement series: (A) a representative autocorrelation function; (B) a representative radius distribution obtained by CONTIN[88] analysis; (C) the synthesis of many radii distribution to show the temporal evolution of particle sizes in the protein droplet. (D) Radial distribution histograms characterizing the hGAPDH oligomerization. A purified sample of hGAPDH was used.

3.3. Protein and protein-probe complex characterization

3.3.1. Circular Dichroism

The conformational and folding states of hGAPDH, either alone or in complex with the CPO-probe, were determined by CD in the far UV region (200-250 nm). The apo-GAPDH spectrum (black line) is indicative of β-sheet secondary structure in their composition (Figure 3.4), displaying a more pronounced minimum at 218 nm.

Moreover, the hGAPDH + CPO probe in ACN at two different incubation times (green line for 15 min of incubation and purple line for 30 min of incubation) showed no significant changes in

the secondary structure of the enzyme. Additionally, we estimated quantitatively the helix content and twist angle distribution for the antiparallel and parallel β-sheets using the BeStSel server[89], showing that there is a considerable percentage of disordered structure (Others) (see Table 3.1). This means that when the CPO-probe (in ACN) binds to hGAPDH, the content of α-helix (~ 9 %), antiparallel β-sheet (~ 34 %), turns (~ 13%) and random coil (~ 43 %) remained unchanged.

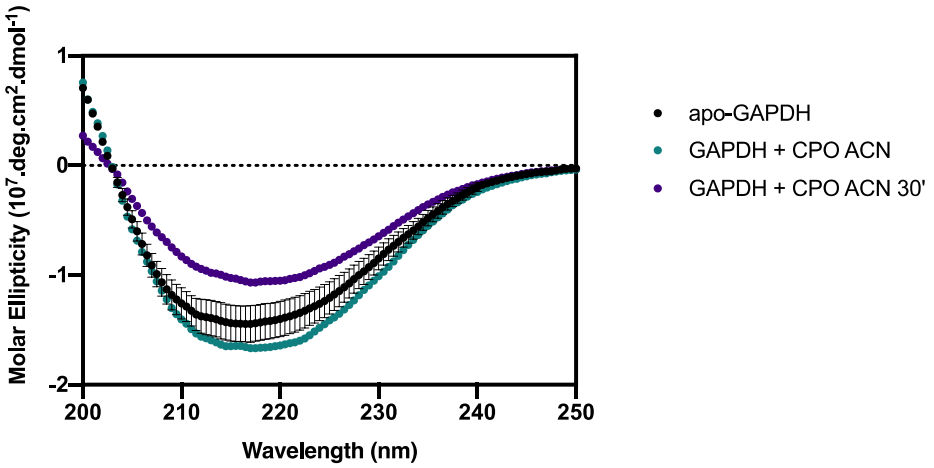

Figure 3.4. Characterization of hGAPDH, either alone or in complex with CPO-probe, by circular dichroism. The apo-hGAPDH displays a curve typical of predominantly β antiparallel secondary structure.

Table 3.1. Estimated secondary structure content (%) using the BeStSel server[89].

	Apo-hGAPDH	hGAPDH – CPO (ACN) 15’	hGAPDH – CPO (ACN) 30’
α – helix	9.1	9.3	4.9
Antiparallel β-sheet	34.3	32.0	35.0
Parallel β-sheet	0.0	0.8	0.0
Turn	13.2	12.5	13.9
Others	43.4	45.3	46.3

3.3.2. Mass spectrometry – LC-MS

Liquid chromatography combined with mass spectrometry (LC–MS) is a valuable and widely recognized technique applied in protein and peptide characterization. MS has become integral part of any proteomics investigation due to its versatility, accuracy, and reproducibility. After expression and purification of the protein, it was necessary to verify the molecular weight of the recombinant GAPDH (predicted to be 38.0 KDa by ProtParam tool)[90]. Thus, an analysis by LC-MS was performed in order to check the molecular weight of the protein (Figure 3.5). The spectra obtained from LC-MS analysis of the recombinant GAPDH shows a predominant peak with a

mass of approximately 38 kDa, confirming the expected molecular weight for this recombinant protein, in agreement with the previous results from SDS-PAGE gel.

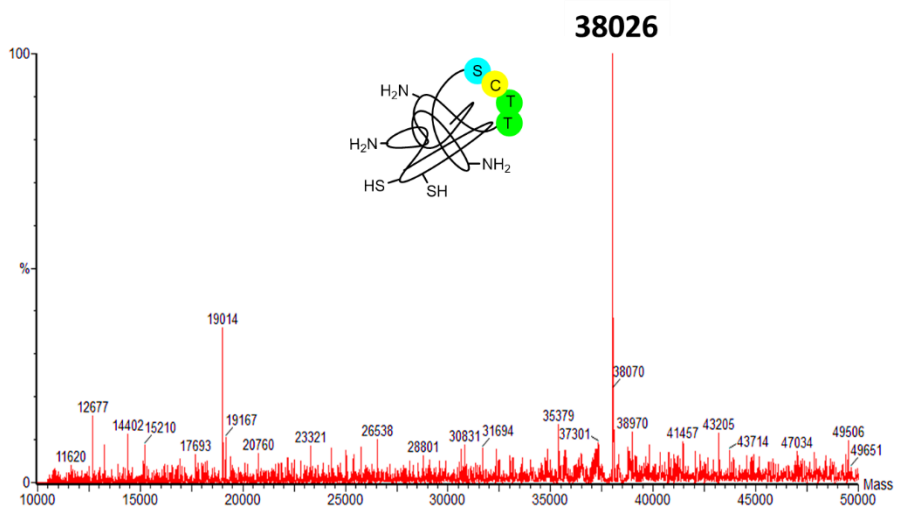

Figure 3.5. Mass spectra of recombinant GAPDH.

Subsequent LC-MS analysis of commercial GAPDH (Mw= 35.9 KDa), either alone or in complex with CPO-probe was performed, in order to probe the conjugation between them (Figure 3.6). These experiments were performed by our colleague Toby Journeaux, at Cambridge University.

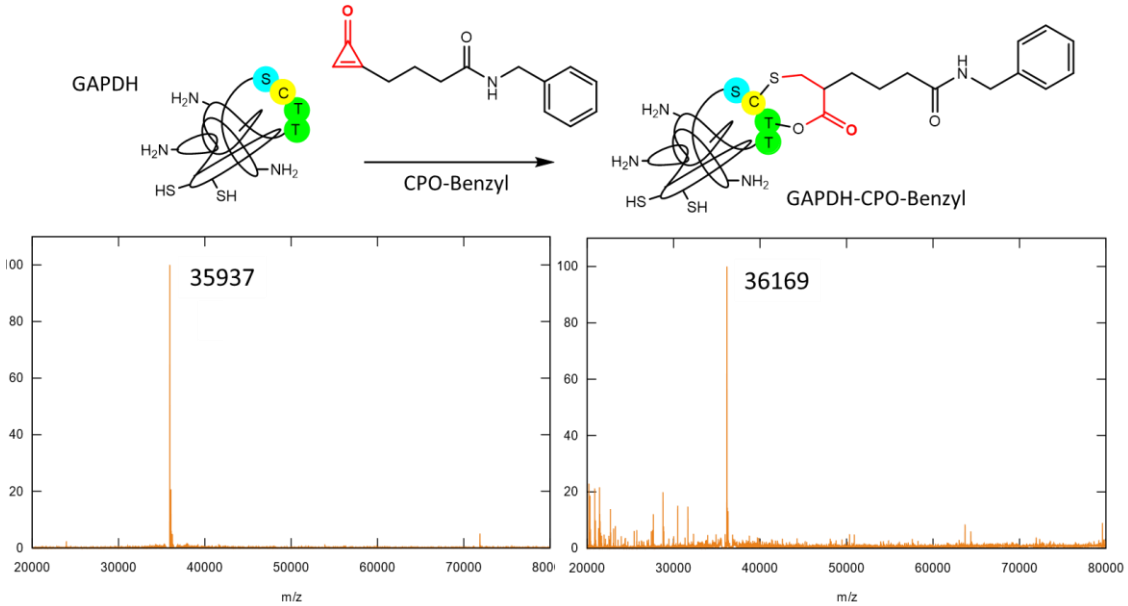

Figure 3.6. Deconvoluted mass spectra of commercial GAPDH before and after conjugation with CPO-Benzyl.

Through the analysis of the spectra obtained we were able to conclude that there was a complete conversion to a singly modified CPO-GAPDH product, after one hour of incubation with CPO-benzyl. This result was confirmed through molecular mass variation, since the variation between the peaks observed in the two spectra is 232 Da (expected $\Delta_{\text{Mass-CPO-Benzyl}}$ = 229 Da).

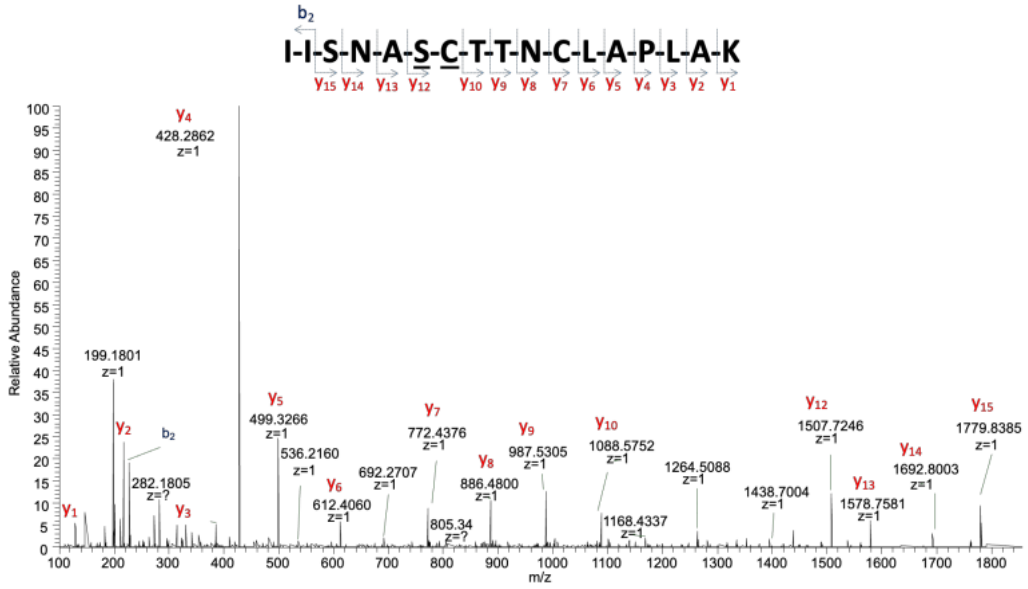

Figure 3.710. MS/MS spectrum of the m/z 1003.5 doubly charged ion of the peptide IISNASCTTNCLAPLAK from hGAPDH containing the CPO-probe modification at the underlined cysteine and serine residues. Work done by Toby Journeaux at Cambridge University.

Figure 3.7 displays the spectra obtained from the modified product's LC-MS/MS analysis, which indicates no fragmentation between the cysteine in the active site (C155) and the neighboring serine residue. This indicates that CPO conjugation takes place between these two residues (Figure 3.8).

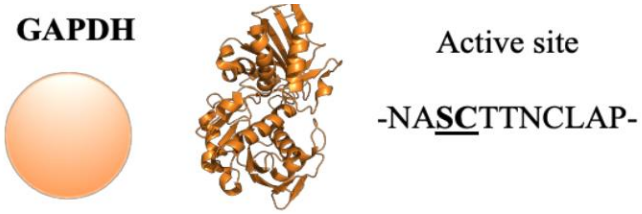

Figure 3.8. GAPDH active site for CPO-probe conjugation.

3.3.3. Differential scanning fluorimetry with the RUBIC Buffer Screen

The RUBIC Buffer Screen[91] is a set of 96 buffers commonly used in structural biology experiments and is designed in such a way that it is possible to discern global stability trends according to pH, salt concentration, buffer type and concentration. These screens are aimed at helping determine protein sample stability, and ideal purification and storage conditions.[91] The DSF data interpretation is based on a melting curve plot describing relative values of the fluorescence intensity observed. The melting temperature (T_m) value of the protein in each state of the RUBIC Buffer Screen needs to be compared with the reference T_m to define a buffer condition that stabilizes the protein.[91] To get a sigmoidal curve that indicates the fraction of the unfolded protein, the fluorescence signal is plotted as a function of temperature. The point of inflection is the melting temperature, at which 50% of the protein is unfolded.

After performing the RUBIC Buffer Screen, it was possible to compare the effect of Tris buffer (the original buffer of the protein after purification) with Phosphate buffer. The different melting points extracted from the melting curves for the more stabilizing and destabilizing conditions are given in the insert. Compared to Tris-HCl (T_m of 52.7 °C), Phosphate Buffer (T_m of 61.5 °C), is shown to be the most stabilizing condition increasing the T_m by 8.8 °C. As a consequence, after this assay, it was decided to make a buffer exchange after purification to 20 mM Sodium Phosphate + 20 mM NaCl, to have a better stabilizing condition for hGAPDH crystallization.

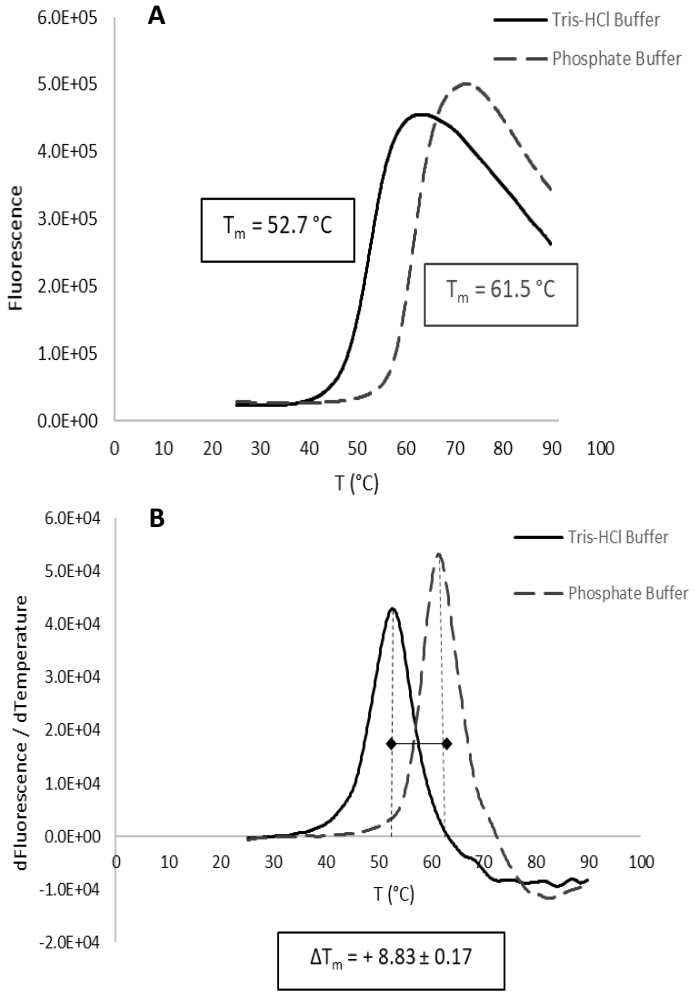

Figure 3.9. Effect of two different buffers on hGAPDH thermal stability. (**A**) Representative melting curve for Tris Buffer and Phosphate Buffer and the respective melting points extracted from the melting curves, (**B**) Thermal melting derivative curves of hGAPDH for Tris Buffer and for Phosphate Buffer. The ΔT_m represents the shift in the melting temperature curve caused by the buffer exchange.

3.3.4. Differential scanning fluorimetry – hGAPDH-CPO probe

Biophysical approaches based on affinity are useful tools for detecting and analyzing molecular interactions between proteins and small molecules. This makes it possible to find and characterize better drug candidates that can modify the activity of a target protein in therapeutic applications.

With the purpose of characterizing the hGAPDH-CPO probe complex, DSF can be valuable to help find the best probe concentration at which hGAPDH maintains its stability.

As a first approach, three different probe concentrations to incubate with hGAPDH (2 mM, 0.5 mM, and 0.25 mM, in Phosphate Buffer) were analyzed by DSF, including hGAPDH apo form for comparison (Figure 3.10).

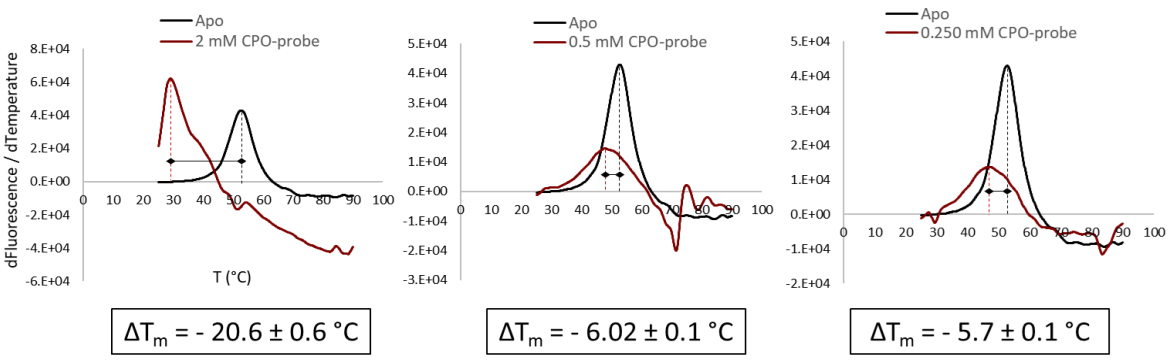

Figure 3.1011. Effect of three different probe concentrations on hGAPDH thermal stability: (A) hGAPDH + 2 mM CPO; (B) hGAPDH + 0.5 mM CPO; (C) hGAPDH + 0.25 mM CPO. The ΔT_m represents the shift in the melting temperature curve caused by the presence of a binder molecule.

From this experiment, we conclude that the presence of the probe destabilizes hGAPDH, as there is a decrease in T_m. However, this decrease is most significant when the probe concentration is higher: ΔT_m for 2 mM and 0.25 mM of CPO-probe is -20.6 °C and -5.7 °C, respectively. These results allow the conclusion that the protein is less destabilized with lower probe concentrations.

3.3.5. GAPDH activity assay with CPO-probe

To further clarify the inhibitory effect of CPO-probe on GAPDH activity, enzyme activity experiments were performed and GAPDH activity was measured.

As a first approach, we tested two different probe concentrations of CPO, 35 µM and 350 µM and the probe was incubated with GAPDH for 15 min, 30 min and 1 h, to see the effects of concentration over time in GAPDH activity. The calculations were done following the manufacturer's protocol [70], using the Equation 3.1., where *B* is the amount (nmole) of NADH generated, *T* is the time reaction incubated in minutes and *V* is the pretreated sample volume (mL) added to each well.

$$\textbf{GAPDH Activity (milliunit/ mL)} = \frac{\text{Bx Sample Dilution Factor}}{\text{T x V}} \qquad \textbf{Eq. (3.1)}$$

Table 3.2. summarizes the calculated values of GAPDH activity, expressed in milliunits per mL, for all the conditions described above (two different concentrations of CPO at three different incubation times), as well as the positive control given by the manufacturers.

Table 3.2. Effect of CPO-probe concentration and different incubation time on GAPDH Activity, expressed in milliunit/mL. The Positive Control was measured at the same time as the other samples.

Concentration of CPO-probe (μM)	GAPDH Activity (milliunit/mL)			Positive Control (milliunit/mL)
	Time of incubation (min)			
	15	30	60	
35	42.22 ± 10.8	40.14 ± 9.63	18.09 ± 0.25	51.24 ± 3.37
350	45.76 ± 11.5	46.63 ± 12.1	21.2 ± 0.41	

If we analyze the graphs depicted in Figure 3.11 A, we can observe that when we incubated GAPDH with the CPO-probe at a concentration of 35 μM, the enzymatic activity of GAPDH progressively decreases across the entire period of incubation with the CPO probe. This behavior of reduction in activity for the GAPDH + CPO-probe sample is more obvious throughout longer incubation times. If we compare GAPDH activity after 60 mins of incubation with the CPO-probe (18.09 milliunits/mL) with GAPDH activity of the positive control (51.24 milliunits/mL), we can see that they are in the same activity range.

It can also be observed that between 15 mins and 30 mins, $\Delta_{\text{GAPDH Activity}}$ is similar (- 2.08 milliunits/mL), which suggests that the CPO-probe has a much greater inhibitory effect the longer the incubation time.

The results obtained with 350 μM CPO-probe concentration did not differ significantly from the assay where 35 μM was used. The maximum $\Delta_{\text{GAPDH Activity}}$ was once again obtained at 60 mins of incubation time (- 30.04 milliunits/mL). While these observations don't allow the establishment of a relationship between the CPO-probe concentration and an inhibitory effect, they are indeed sufficient to determine that the incubation time of the CPO-probe with the enzyme does have an inhibitory impact on GAPDH activity.

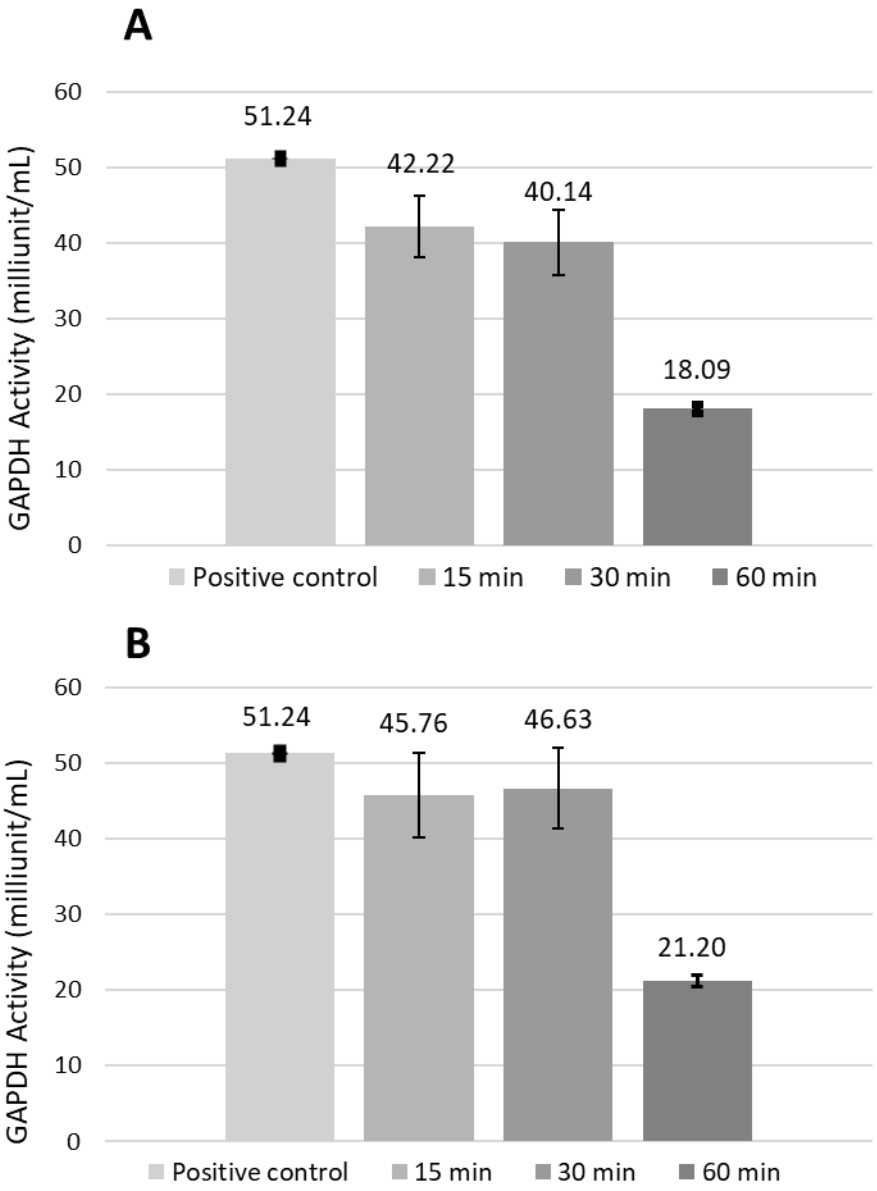

Figure 3.11. Graphic illustration of the effect of CPO-probe on GAPDH Activity; (A) 35 µM of CPO-probe, incubated with GAPDH at 15, 30 and 60 min; (B) 350 µM of CPO-probe, with GAPDH at 15, 30 and 60 min of incubation.

As a second approach, we tested DMSO effect (present in CPO-probe as main solvent) on GAPDH activity and its effect as a function of incubation time. Two different incubation times were analyzed, 30 min and 60 min in a fresh sample of purified GAPDH, and 60 min of incubation with an old sample of GAPDH. The probe concentration chosen was 35 µM and a new positive control was calculated for this assay. Calculations were done following the manufacturer's protocol [70] by using the Equation 3.1.

Table 3.3. Effect of DMSO on GAPDH Activity, in different incubation times with CPO-probe, expressed in milliunit/mL. The Positive Control was measured at the same time as the other samples.

Sample	Time of incubation (min)	GAPDH Activity (milliunit/mL) + DMSO	GAPDH Activity (milliunit/mL) + CPO	Positive Control (milliunit/mL)
Fresh	30	4.15 ± 0.37	2.48 ± 0.15	18.23 ± 2.4
	60	8.73 ± 0.03	3.57 ± 0.4	
Old	60	5.91 ± 0.88	6.24 ± 0.88	

The graph depicted on Figure 3.12, clearly illustrates the effect of DMSO on GAPDH activity. First, if we compare the values of GAPDH activity under the 3 conditions (CPO-30 min fresh, CPO-60 min fresh and CPO-60 min old). These results are in agreement with the results obtained previously, demonstrating the inhibitory effect of CPO-probe on enzyme activity.

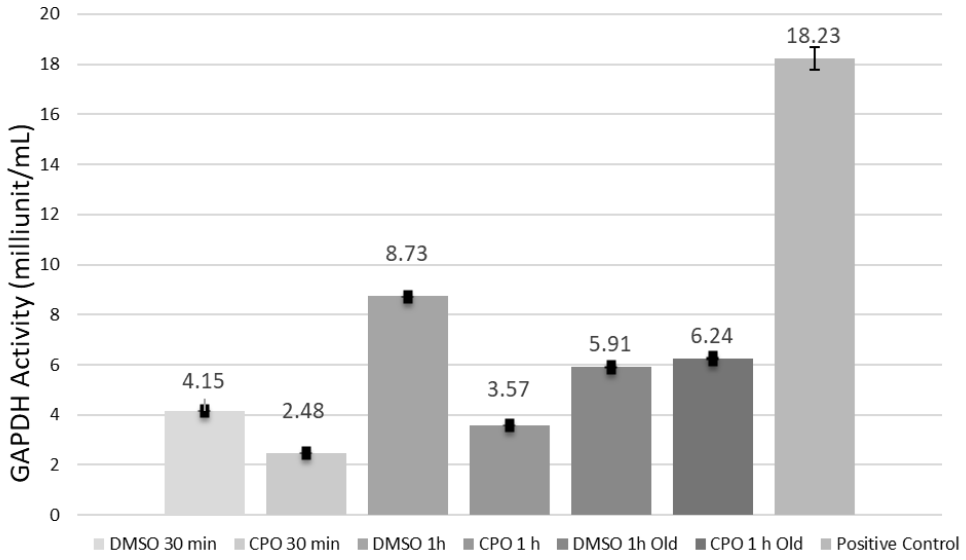

Figure 3.12. Graphic illustration of the effect of DMSO on GAPDH Activity (35 µM CPO-probe). Different conditions are represented: DMSO and CPO with an incubation time of 30 min and 60 min on a fresh sample of purified GAPDH and DMSO and CPO effect on an old sample of GAPDH, at 60 min of incubation.

Next, the variations in GAPDH activity between the DMSO and the probe were analyzed. From the graph, it can be seen that when a fresh sample of GAPDH is incubated with DMSO or CPO for 30 min, there is a variation in activity of 1.67 milliunit/mL. When the incubation time is doubles, the variation increases to 5.16 milliunit/mL. However, if DMSO and CPO-probe are incubated with an older sample of GAPDH, no differences are observed.

With this analysis it can be concluded that DMSO has in fact an influence in GAPDH activity. Therefore, even though the CPO-probe has a substantial inhibitory power, it must always be assumed that some of that inhibitory power comes from the DMSO. Another of the conclusions drawn from this experiment is that, when an older purified GAPDH sample is used, it completely loses its activity, becoming invalid for long-term uses.

3.4. Protein structure determination by X-ray crystallography

3.4.1. GAPDH and GAPDH-CPO complex crystallization assays

To date, many structures of human GAPDH have been deposited on PDB and one of the main thesis objectives was, in addition to obtaining the crystallographic structure of hGAPDH in complex with the CPO-probe, to obtain a structure of the apo form of hGAPDH with higher resolution than those already deposited by Ismail & Park in 2005 (PDB code 1znq, 2.5 Å)[35] and Jenkins and Tanner in 2006 (PDB code 1u8f, 1.75 Å[54]).

An initial crystallization trial was set up with purified hGAPDH alone using the conditions reported in White *et al.*, in 2015[71]. From this first trial, the crystals presented on Figure S1 were obtained. However, after crystal optimization, X-ray diffraction tests showed that these were in fact salt crystals.

Therefore, it was decided to carry out crystallization screens using two commercial crystallization screen kits (Screen 1+2 and BSC Screen) on a crystallization robot. Crystallization trials were set up with hGAPDH alone or pre-incubated with CPO-probe in ACN and in DMSO. The most promising results are shown in Figure S2. To make sure that the crystals obtained were in fact protein crystals, the Izit crystal dye assay[92] was performed. Izit is a small molecule dye which will fill the solvent channels in protein crystals, coloring the crystals blue. Salt crystals do not possess these large solvent channels, therefore, Izit cannot enter salt crystals, thus leaving them colorless. In this way it was possible to verify which conditions gave rise to salt crystals, and thus discard them, and which conditions formed protein crystals and thus optimize them (Figure S3).

After several attempts to crystallize the apo form of GAPDH by optimizing the most promising screen conditions, crystals suitable for X-ray diffraction were obtained. The condition corresponded to 1.6 M of Sodium Citrate, pH 7 and the crystals appeared after 7 days (Figure 3.13). The apo-GAPDH crystal was cryoprotected with a crystallization solution to which 25% v/v glycerol were added, flash-cooled in liquid nitrogen, and sent to the ESRF for X-ray data collection.

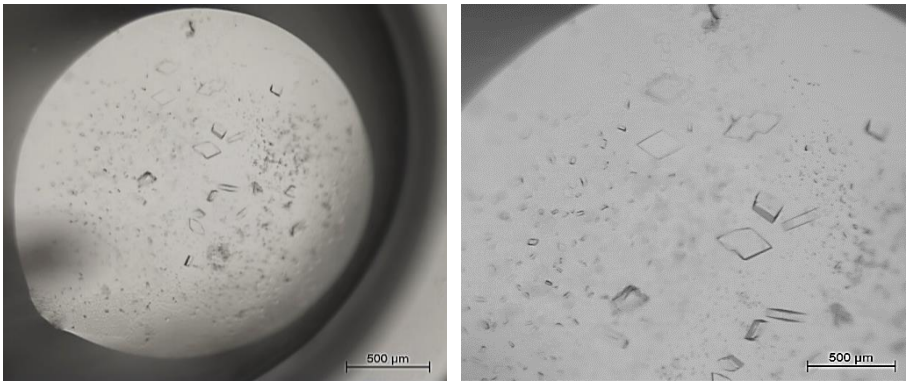

Figure 3.13. Optical microscopy image of apo-hGAPDH crystals. Sitting-drop vapor diffusion method; [GAPDH] = 263.2 µM; **Condition:** 1.6 M Sodium Citrate, pH 7.

After obtaining crystals of the apo form of GAPDH, work continued to crystallize the GAPDH-CPO complex, and crystallization trials with GAPDH pre-incubated with CPO-probe in DMSO were set up using a crystallization robot with Screen 1+2 and BSC Screen. The most promising results are shown in Figure S4. The best conditions were optimized and this time, 1 hit condition was obtained, consisting of 0.2 mM Tris pH 8.5, 21% PEG 3350. The crystals appeared after 4 days and reached their full size in 1 week (Figure 3.14). The GAPDH-CPO crystals wwere cryoprotected with a crystallization solution containing an additional 30% v/v glycerol, flash-cooled in liquid nitrogen, and sent to the ESRF for X-ray data collection.

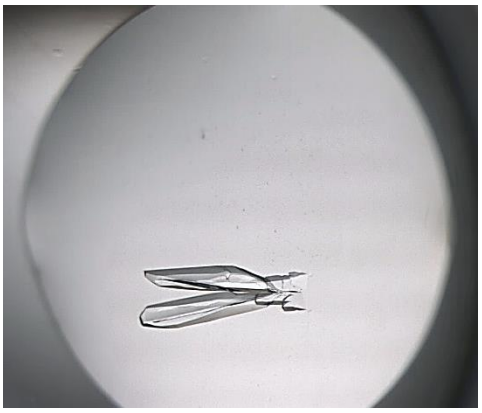

Figure 3.14. Optical microscopy image of hGAPDH crystals in complex with CPO-probe in DMSO.

Despite the fact that all the crystallization trials were prepared with hGAPDH previously incubated with the CPO-probe, the obtained crystal structure was in fact of the apo form of the protein, without the presence of cofactor NADH. The diffraction data were collected to a resolution of 1.7 Å, however, the crystal structure was not refined, due to lack of time.

3.4.2. Overall structure of hGAPDH in complex with co-factor NADH

The X-ray crystal structure of hGAPDH in complex with NADH has been determined to 2.11 Å resolution (Figure 3.15).

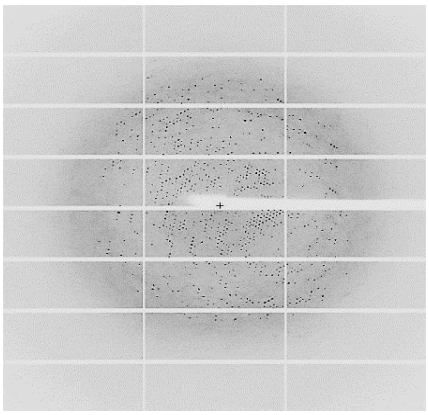

Figure 3.15. Diffraction image of a hGAPDH crystal. The black spots on the image are the result of the cooperative scattering (diffraction) from the electrons of all atoms contained in the crystal. Analysis of the pattern yields information about the structure of the protein.

hGAPDH crystallized in space group *I* 222 with three independent molecules in asymmetric unit, are arranged in two different tetramers, generated by crystallographic symmetry: the first is formed by two monomers related by a 2-fold non-crystallographic symmetry axis and their crystallographic symmetry equivalents; the second is formed by a single monomer sitting on a site with 222 crystallographic symmetry (Figure 3.16). Chain A includes residues from 6 to 338 and chain B and chain C includes the residues from -1 to 338. Residues -1 and 0 in chain B and C belong to the TEV cleavage site. The superposition of the three monomers (A, B and C) in the asymmetric unit of hGAPDH (Figure 3.17) demonstrated that the three chains have a nearly identical fold. The r.m.s.d. determined between the three chains was very low, 0.34 Å for the chain B to chain A superposition, and 0.25 Å for the chain C to chain A superposition, with 333 aligned residues in both alignments.

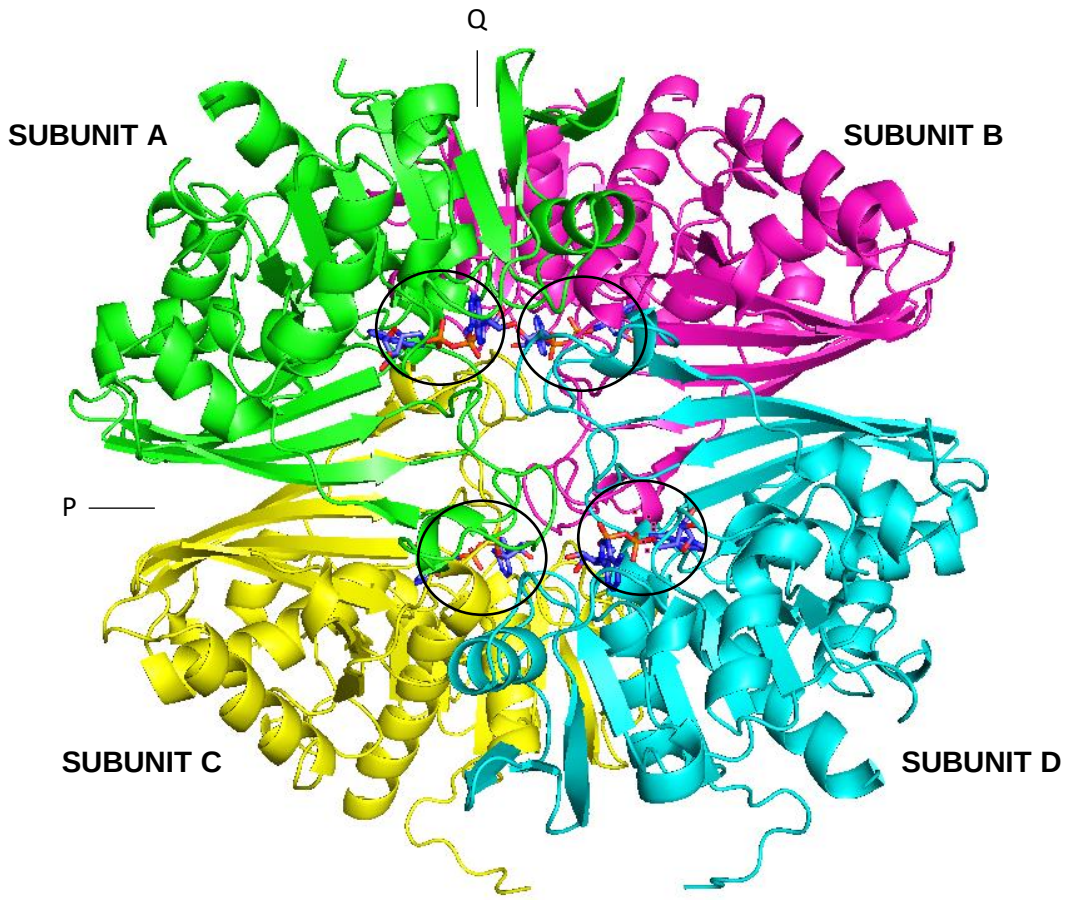

Figure 3.16. Overall view of the hGAPDH homotetramer. Subunits A, B, C and D are shown in green, magenta, yellow and cyan respectively. Each monomer is bound to an NADH molecule (dark blue, highlighted in figure with black circles). This figure was produced with *PyMOL*.

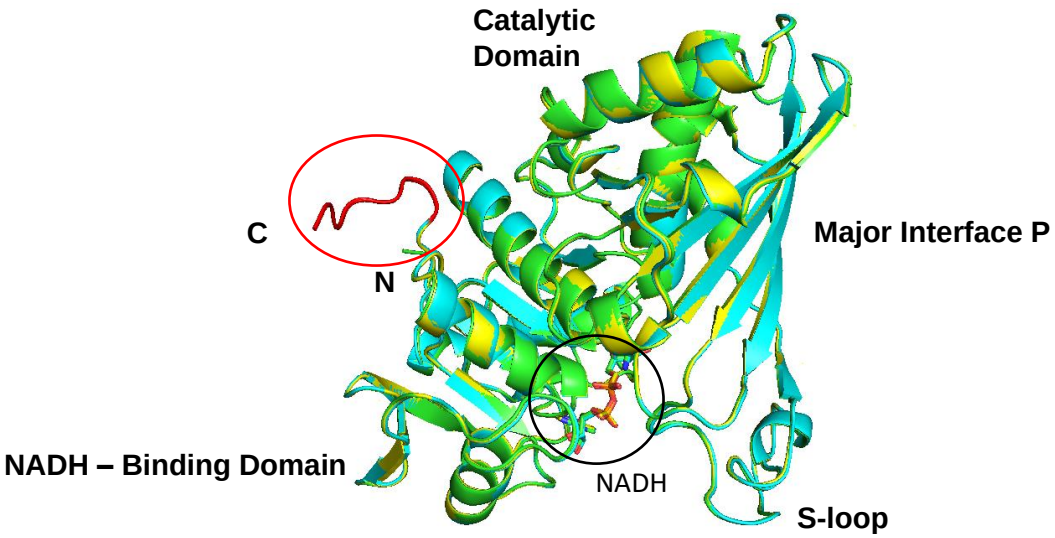

Figure 3.17. Superposition of the three independent hGAPDH chains in the crystal structure. Chain A is colored green, chain B in cyan and chain C in yellow. The co-factor NADH is marked in the figure (black circle). Residues shown in red correspond to residues -1 to 5, only present in the B and C chains (red circle). This figure was produced with *PyMOL*.

hGAPDH occurs as a homotetramer (~144 kDa), composed of four identical subunits A, B, C and D, each approximately 37 kDa in molecular mass, and each harboring an active site. (Figure 3.16). Each monomer contains two major domains, the NADH-binding domain which includes most of the polypeptide chain, and the catalytic glyceraldehyde-3-phosphate (G3P)-binding domain which contains the C-terminal helix (Figure 3.17).

The NADH-binding domain, comprising amino acids 1-155, displays the characteristic Rossmann fold [93], with a central parallel β-sheet flanked on both sides by α-helices (αβα), with catalytic cysteine (Cys155). This area is responsible for binding dinucleotides. The catalytic domain of GAPDH (residues 156-338) folds into an eight-stranded antiparallel β-sheet and four α-helices. This area is responsible for the specific catalytic activities of the reaction sequence, and substrate specificity.

The active center of the enzyme is located in a large slot between the catalytic domain and the coenzyme binding domain. This area is a large recess, in order to accommodate G3P with a phosphate ion and the coenzyme NADH. The amino acids which are directly involved in binding NADH and G3P are Cys155, His182, Asp38, Arg237, Thr185, Thr214, Tyr317, and Tyr323.[94]

Each of the four subunits of human GAPDH contains 3 cysteine residues (Cys155, Cys159, and Cys250), two of them located in the active site of the enzyme and easily subjected to oxidation and other modifications. The electron density shown in Figure 18 indicates that the active-site cysteine (Cys155) residue is oxidized to sulfenic acid or S-oxyl Cys.

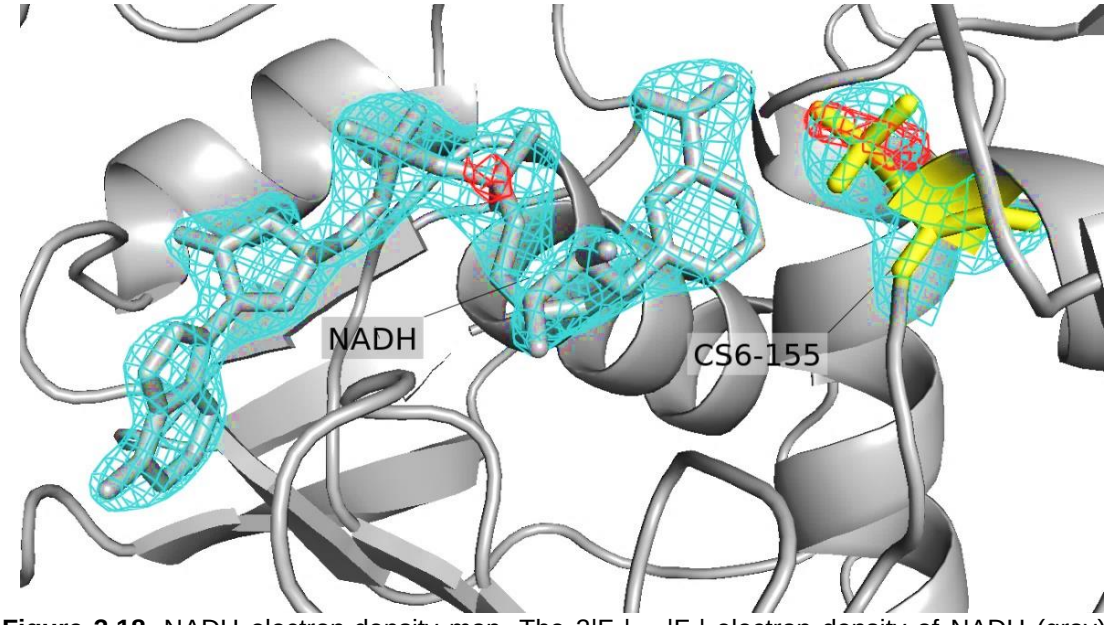

Figure 3.18. NADH electron-density map. The $2|F_o| - |F_c|$ electron density of NADH (gray) is contoured at 1.5σ (cyan) and the $|F_o| - |F_c|$ electron density is contoured at 3.0σ (green) and -3.0σ (red). The oxidized active-site cysteine 155 is represented in yellow. The figure was produced with PyMOL.

Despite many attempts, the structure of hGAPDH with the CPO-probe has not been determined. There is a possibility that the produced hGAPDH does not bind 100% to the CPO-probe, thus hindering the crystallization of the complex. Future efforts will be made to obtain a crystal from the complex, and thus determine its structure.

3.5. Cell Biology

3.5.1. Intracellular GAPDH activity assay and inhibition with the CPO-probe

To investigate whether the CPO-probe can effectively inhibit cancer proliferation, several cell lines were selected in which GAPDH was overexpressed. MCF7 cells, HepG2 cells and A549 cells were incubated with CPO-probe, and GAPDH enzyme activity was measured after 48 h of treatment. CPO-probe was found to inhibit the proliferation in most cancer cell lines, with the greatest effect observed for HepG2 cells, where an IC_{50} value of 35 µM was recorded (Figure 3.19). CPO-probe had no effect on the proliferation of triple-negative breast cancer MDA-MB-231 cells and minimal effect on HEK293T cells, where a larger standard deviation was observed (Figure 3.20). This behavior was likely caused by the semi-adherent characteristics of HEK293T cells.

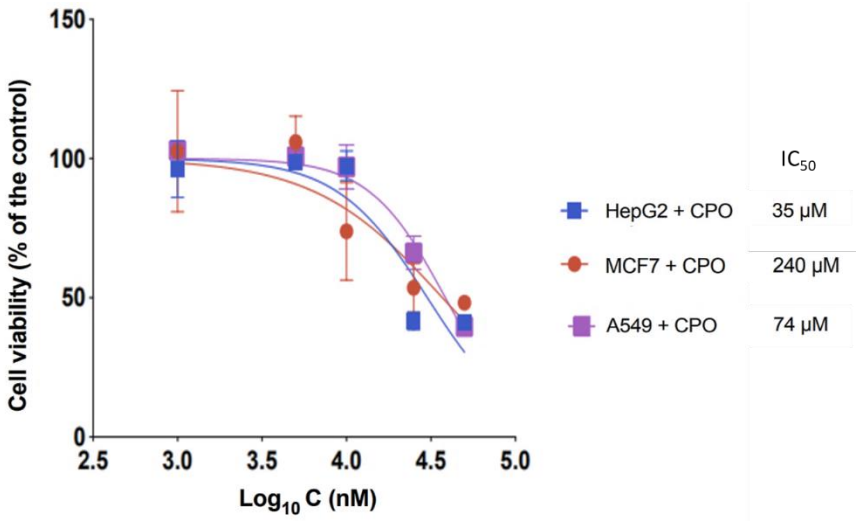

Figure 3.19. Cell viability analysis, showing the effect of CPO reagent on the proliferation of a range of cancer cell lines (MCF7 cells, A549 cells and HepG2 cells). IC_{50} values were calculating using variable slope fitting function.

Since CPO-probe had the greatest effect on HepG2 cell proliferation, the mechanism of action of CPO-probe in HepG2 cells will be investigated in the future. This will involve performing GAPDH activity assays, as well as glucose and lactate assays on CPO and KA treated cells. These experiments will verify whether the reduced cell proliferation was caused by the inhibition of GAPDH by CPO and will provide a comparison to the best-known inhibitor.[8] Once complete,

this work hopes to highlight both the utility of CPO as a potential cancer therapeutic but also demonstrate the capacity of CPO to selectively modify and probe native proteins.

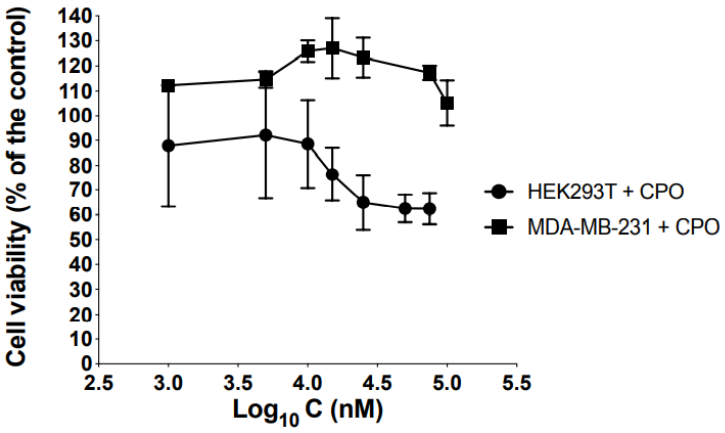

Figure 3.20. Cell viability analysis of HEK293T and MDA-MD-231 cell lines, highlighting the effect of CPO reagent on cell proliferation.

Chapter 4

4. Conclusions and future perspectives

This master thesis plan was aimed at obtaining a functional and structural characterization of the glyceraldehyde 3-phosphate dehydrogenase in complex with a CPO-probe to understand the molecular basis of the complex and to explore CPO-probes as selective inhibitors.

To summarize, in this work we described a method for producing a widely known protein, the human GAPDH, bearing in the N-terminal region a Strep-tag II fused to a TEV cleavage site, using the common bacterial expression system *E. coli* BL21(DE3). Purification gave rise to a satisfactory yield of 7.72 mg of protein per liter of culture. The purified samples were submitted to biochemical and biophysical characterization by SDS-PAGE, Western-Blot, Dynamic Light Scattering, and Circular Dichroism, allowing the protein quality assessment before crystallization.

DLS measurements were performed to evaluate the dispersion level of the particles in solution and their aggregation state, and the results confirmed that the purified sample was monodisperse and stable in solution. The conformational and folding states of hGAPDH, either alone or in complex with CPO-probe, were determined by CD and showed the protein to have a predominantly β-antiparallel secondary structure. The hGAPDH in complex with the CPO-probe in ACN was like its native fold state. The spectra obtained from LC-MS analysis for the recombinant GAPDH showed a predominant peak with a mass of approximately 38 kDa, confirming the expected molecular weight for this protein, in agreement with the SDS-PAGE gel results. Subsequent LC-MS analysis of commercial GAPDH, either alone or in complex with CPO-probe proved, through molecular mass variation, that the CPO-probe binds to hGAPDH. A DSF assay with the RUBIC Buffer Screen demonstrated the best buffer conditions for hGAPDH to be phosphate buffer at pH 7.4. By DSF experiments it was also possible to conclude that the presence of the probe destabilizes hGAPDH, as there is a decrease in T_m. Furthermore, this decrease is most significant for higher probe concentrations.

Several cell biology assays, such as measuring the intracellular GAPDH activity in cancer cells, were also carried out with the aim of verifying cell viability upon CPO-probe incubation. The CPO was found to inhibit the proliferation of most cancer cell lines, with the greatest effect observed for HepG2 cells, with an IC_{50} value of 35 µM.

The final stage of this thesis plan included crystallization trials with the hGAPDH alone and in complex with the CPO-probe. The X-ray crystal structure of hGAPDH in complex with NADH has been determined to 2.11 Å resolution. hGAPDH crystallized in space group *I* 222 with three independent molecules in the asymmetric unit. The biological unit of hGAPDH is a homotetramer (~ 144 kDa), composed of four identical subunits.

Despite many attempts, the structure of hGAPDH with the CPO-probe has not been determined. Soaking of hGAPDH crystals did not yield the crystal structure of the complex, and incubation of GAPDH with the CPO-probe prior to crystallization very likely resulted in a heterogeneous mixture in which only the part of the hGAPDH molecules do not contain the bound CPO-probe crystallized. These results were corroborated by a new LC-MS analysis, where we verified that the recombinant hGAPDH did not completely bind to the CPO-probe (results not shown).

With regards to further experiments, new efforts will be made to obtain the 3D structure of the complex. To achieve this, our efforts will be focused on obtaining a 100% binding between hGAPDH and the CPO-probe by modifying the probe solvent, optimizing the CPO-hGAPDH incubation protocol or by cleaving the tags present at the N-terminal. Moreover, we also plan to perform glucose and lactate assays with cells treated with CPO and koningic acid (KA). These experiments will indicate whether the reduced cell proliferation is caused by the modulation of GAPDH by the CPO and if this CPO-probe could serve as GAPDH inhibitor.

Hopefully, this work will highlight both the utility of CPO-probes as potential cancer therapeutics through GAPDH inactivation, but also demonstrate the capacity of CPO to selectively modify and probe native proteins. CPO reacts selectively and almost exclusively with native proteins containing internal cysteines adjacent to a serine or threonine residue. This discovery prompted investigations into native protein labelling as well as the development of a CPO motif which can theoretically be incorporated into any protein. The modification of the native protein GAPDH was of particular interest due to the role it plays in cancer cell proliferation.

The determination of the 3D-dimensional structure of the hGAPDH in complex with this CPO-probe will play a key role for directing further studies aiming to develop therapeutically active molecules and might represent a major advance in future cancer therapy.

Bibliography

(1) Cancer https://www.who.int/health-topics/cancer#tab=tab_1 (accessed Oct 26, 2021).

(2) Brennan, K.; Offiah, G.; McSherry, E. A.; Hopkins, A. M. Tight Junctions: A Barrier to the Initiation and Progression of Breast Cancer? *J. Biomed. Biotechnol.* **2010**, *2010*. https://doi.org/10.1155/2010/460607.

(3) de Martel, C.; Georges, D.; Bray, F.; Ferlay, J.; Clifford, G. M. Global Burden of Cancer Attributable to Infections in 2018: A Worldwide Incidence Analysis. *Lancet Glob. Heal.* **2020**, *8* (2), e180–e190. https://doi.org/10.1016/S2214-109X(19)30488-7.

(4) GLOBOCAN. The Global Cancer Observatory - All Cancers. *Int. Agency Res. Cancer - WHO* **2020**, *419*, 199–200.

(5) Kreeger, P. K.; Lauffenburger, D. A. Cancer Systems Biology: A Network Modeling Perspective. *Carcinogenesis* **2009**, *31* (1), 2–8. https://doi.org/10.1093/carcin/bgp261.

(6) What Is Cancer? - National Cancer Institute https://www.cancer.gov/about-cancer/understanding/what-is-cancer (accessed Oct 27, 2021).

(7) Kobayashi, Y.; Banno, K.; Kunitomi, H.; Takahashi, T.; Takeda, T.; Nakamura, K.; Tsuji, K.; Tominaga, E.; Aoki, D. Warburg Effect in Gynecologic Cancers. *J. Obstet. Gynaecol. Res.* **2019**, *45* (3), 542–548. https://doi.org/10.1111/jog.13867.

(8) Li, T.; Tan, X.; Yang, R.; Miao, Y.; Zhang, M.; Xi, Y.; Guo, R.; Zheng, M.; Li, B. Discovery of Novel Glyceraldehyde-3-Phosphate Dehydrogenase Inhibitor via Docking-Based Virtual Screening. *Bioorg. Chem.* **2020**, *96* (September 2019), 103620. https://doi.org/10.1016/j.bioorg.2020.103620.

(9) O. Warburg. On the Origin of Cancer Cells. *Science (80-.).* **1956**, *123* (3191), 309–314.

(10) Hsu, P. P.; Sabatini, D. M. Cancer Cell Metabolism: Warburg and Beyond. *Cell* **2008**, *134* (5), 703–707. https://doi.org/10.1016/j.cell.2008.08.021.

(11) Devic, S. Warburg Effect - a Consequence or the Cause of Carcinogenesis? *J. Cancer* **2016**, *7* (7), 817–822. https://doi.org/10.7150/jca.14274.

(12) El Sayed, S. M.; Mahmoud, A. A.; El Sawy, S. A.; Abdelaal, E. A.; Fouad, A. M.; Yousif, R. S.; Hashim, M. S.; Hemdan, S. B.; Kadry, Z. M.; Abdelmoaty, M. A.; Gabr, A. G.; Omran, F. M.; Nabo, M. M. H.; Ahmed, N. S. Warburg Effect Increases Steady-State ROS Condition in Cancer Cells through Decreasing Their Antioxidant Capacities (Anticancer Effects of 3-Bromopyruvate through Antagonizing Warburg Effect). *Med. Hypotheses* **2013**, *81* (5), 866–870. https://doi.org/10.1016/j.mehy.2013.08.024.

(13) Brahimi-Horn, M. C.; Chiche, J.; Pouysségur, J. Hypoxia Signalling Controls Metabolic Demand. *Curr. Opin. Cell Biol.* **2007**, *19* (2), 223–229. https://doi.org/10.1016/j.ceb.2007.02.003.

(14) Vazquez, A.; Liu, J.; Zhou, Y.; Oltvai, Z. N. Catabolic Efficiency of Aerobic Glycolysis: The Warburg Effect Revisited. *BMC Syst. Biol.* **2010**, *4*. https://doi.org/10.1186/1752-0509-4-58.

(15) Gatenby, R. A.; Gillies, R. J. Why Do Cancers Have High Aerobic Glycolysis? *Nat. Rev. Cancer* **2004**, *4* (11), 891–899. https://doi.org/10.1038/nrc1478.

(16) Seidler, N. W. *Basic Biology of GAPDH*; 2013; Vol. 985. https://doi.org/10.1007/978-94-007-4716-6_1.

(17) Nicholls, C.; Li, H.; Liu, J. P. GAPDH: A Common Enzyme with Uncommon Functions. *Clin. Exp. Pharmacol. Physiol.* **2012**, *39* (8), 674–679. https://doi.org/10.1111/j.1440-1681.2011.05599.x.

(18) Tristan, C.; Shahani, N.; Sedlak, T. W.; Sawa, A. The Diverse Functions of GAPDH: Views from Different Subcellular Compartments. *Cell. Signal.* **2011**, *23* (2), 317–323. https://doi.org/10.1016/j.cellsig.2010.08.003.

(19) Cerella, C.; Dicato, M.; Diederich, M. Modulatory Roles of Glycolytic Enzymes in Cell Death. *Biochem. Pharmacol.* **2014**, *92* (1), 22–30. https://doi.org/10.1016/j.bcp.2014.07.005.

(20) Colell, A.; Green, D. R.; Ricci, J. E. Novel Roles for GAPDH in Cell Death and Carcinogenesis. *Cell Death Differ.* **2009**, *16* (12), 1573–1581. https://doi.org/10.1038/cdd.2009.137.

(21) SCHUPPE-KOISTINEN, I.; MOLDÉUS, P.; BERGMAN, T.; COTGREAVE, I. A. S-Thiolation of Human Endothelial Cell Glyceraldehyde-3-phosphate Dehydrogenase after Hydrogen Peroxide Treatment. *Eur. J. Biochem.* **1994**, *221* (3), 1033–1037. https://doi.org/10.1111/j.1432-1033.1994.tb18821.x.

(22) Tarze, A.; Deniaud, A.; Le Bras, M.; Maillier, E.; Molle, D.; Larochette, N.; Zamzami, N.; Jan, G.; Kroemer, G.; Brenner, C. GAPDH, a Novel Regulator of the pro-Apoptotic Mitochondrial Membrane Permeabilization. *Oncogene* **2007**, *26* (18), 2606–2620. https://doi.org/10.1038/sj.onc.1210074.

(23) Tatton, N. A. Increased Caspase 3 and Bax Immunoreactivity Accompany Nuclear GAPDH Translocation and Neuronal Apoptosis in Parkinson's Disease. *Exp. Neurol.* **2000**, *166* (1), 29–43. https://doi.org/10.1006/exnr.2000.7489.

(24) Tsuchiya, K.; Tajima, H.; Yamada, M.; Takahashi, H.; Kuwae, T.; Sunaga, K.; Katsube, N.; Ishitani, R. Disclosure of a Pro-Apoptotic Glyceraldehyde-3-Phosphate Dehydrogenase Promoter: Anti-Dementia Drugs Depress Its Activation in Apoptosis. *Life Sci.* **2004**, *74* (26), 3245–3258. https://doi.org/10.1016/j.lfs.2003.11.029.

(25) Burke, James; Enghild, Jan; Martin, M. Huntingtin and DRPLA Proteins Selectively Interact with the Enzyme GAPDH. **1996**.

(26) Cumming, R. C.; Schubert, D. Amyloid-β Induces Disulfide Bonding and Aggregation of GAPDH in Alzheimer's Disease. *FASEB J.* **2005**, *19* (14), 2060–2062. https://doi.org/10.1096/fj.05-4195fje.

(27) Kunjithapatham, R.; Ganapathy-Kanniappan, S. GAPDH with NAD+-Binding Site Mutation Competitively Inhibits the Wild-Type and Affects Glucose Metabolism in Cancer. *Biochim. Biophys. Acta - Gen. Subj.* **2018**, *1862* (12), 2555–2563. https://doi.org/10.1016/j.bbagen.2018.08.001.

(28) Moniot, S.; Bruno, S.; Vonrhein, C.; Didierjean, C.; Boschi-Muller, S.; Vas, M.; Bricogne, G.; Branlant, G.; Mozzarelli, A.; Corbier, C. Trapping of the Thioacylglyceraldehyde-3-Phosphate Dehydrogenase Intermediate from Bacillus Stearothermophilus: Direct Evidence for a Flip-Flop Mechanism. *J. Biol. Chem.* **2008**, *283* (31), 21693–21702. https://doi.org/10.1074/jbc.M802286200.

(29) Guo, C.; Liu, S.; Sun, M. Z. Novel Insight into the Role of GAPDH Playing in Tumor. *Clin. Transl. Oncol.* **2013**, *15* (3), 167–172. https://doi.org/10.1007/s12094-012-0924-x.

(30) Phadke, M. S.; Krynetskaia, N. F.; Mishra, A. K.; Krynetskiy, E. Glyceraldehyde 3-Phosphate Dehydrogenase Depletion Induces Cell Cycle Arrest and Resistance to Antimetabolites in Human Carcinoma Cell Lines. *J. Pharmacol. Exp. Ther.* **2009**, *331* (1), 77–86. https://doi.org/10.1124/jpet.109.155671.

(31) Rahier, N. J.; Molinier, N.; Long, C.; Deshmukh, S. K.; Kate, A. S.; Ranadive, P.; Verekar, S. A.; Jiotode, M.; Lavhale, R. R.; Tokdar, P.; Balakrishnan, A.; Meignan, S.; Robichon, C.; Gomes, B.; Aussagues, Y.; Samson, A.; Sautel, F.; Bailly, C. Anticancer Activity of Koningic Acid and Semisynthetic Derivatives. *Bioorganic Med. Chem.* **2015**, *23* (13), 3712–3721. https://doi.org/10.1016/j.bmc.2015.04.004.

(32) Endo, A.; Hasumi, K.; Sakai, K.; Kanbe, T. Specific Inhibition of Glyceraldehyde-3-Phosphate Dehydrogenase by Koningic Acid (Heptelidic Acid). *J. Antibiot. (Tokyo).* **1985**, *38* (7), 920–925. https://doi.org/10.7164/antibiotics.38.920.

(33) Nakazawa, M.; Uehara, T.; Nomura, Y. Koningic Acid (a Potent Glyceraldehyde-3-Phosphate Dehydrogenase Inhibitor)-Induced Fragmentation and Condensation of DNA in NG108-15 Cells. *J. Neurochem.* **1997**, *68* (6), 2493–2499.

https://doi.org/10.1046/j.1471-4159.1997.68062493.x.

(34) Prati, F.; Bergamini, C.; Molina, M. T.; Falchi, F. 2-Phenoxy-1 , 4-Naphthoquinones : From a Multitarget Antitrypanosomal to a Potential Antitumor Profile. 1–44.

(35) Ismail, S. A.; Park, H. W. Structural Analysis of Human Liver Glyceraldehyde-3-Phosphate Dehydrogenase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, *61* (11), 1508–1513. https://doi.org/10.1107/S0907444905026740.

(36) Stephanopoulos, N.; Francis, M. B. Choosing an Effective Protein Bioconjugation Strategy. *Nat. Chem. Biol.* **2011**, *7* (12), 876–884. https://doi.org/10.1038/nchembio.720.

(37) Kwan, T. O. C.; Reis, R.; Siligardi, G.; Hussain, R.; Cheruvara, H.; Moraes, I. Selection of Biophysical Methods for Characterisation of Membrane Proteins. **2019**.

(38) Hoyt, E. A.; Cal, P. M. S. D.; Oliveira, B. L.; Bernardes, G. J. L. Contemporary Approaches to Site-Selective Protein Modification. *Nat. Rev. Chem.* **2019**, *3* (3), 147–171. https://doi.org/10.1038/s41570-019-0079-1.

(39) De Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural Amino Acids for Site-Specific Protein Conjugation. *Bioconjug. Chem.* **2009**, *20* (7), 1281–1295. https://doi.org/10.1021/bc800294a.

(40) Li, B.; Fooksa, M.; Heinze, S.; Meiler, J. Finding the Needle in the Haystack: Towards Solving the Protein-Folding Problem Computationally. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53* (1), 1–28. https://doi.org/10.1080/10409238.2017.1380596.

(41) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms5740.

(42) Dill, K. A.; Maccallum, J. L.; Folding, P. The Protein-Folding Problem , 50 Years On. **2012**, No. november, 1042–1047.

(43) Griffiths, R. C.; Smith, F. R.; Long, J. E.; Williams, H. E. L.; Layfield, R.; Mitchell, N. J. Site-Selective Modification of Peptides and Proteins via Interception of Free-Radical-Mediated Dechalcogenation. *Angew. Chemie - Int. Ed.* **2020**, *59* (52), 23659–23667. https://doi.org/10.1002/anie.202006260.

(44) Reddy, N. C.; Kumar, M.; Molla, R.; Rai, V. Chemical Methods for Modification of Proteins. *Org. Biomol. Chem.* **2020**, *18* (25), 4669–4691. https://doi.org/10.1039/d0ob00857e.

(45) Szijj, P. A.; Kostadinova, K. A.; Spears, R. J.; Chudasama, V. Tyrosine Bioconjugation-an Emergent Alternative. *Org. Biomol. Chem.* **2020**, *18* (44), 9018–9028. https://doi.org/10.1039/d0ob01912g.

(46) Shadish, J. A.; DeForest, C. A. Site-Selective Protein Modification: From Functionalized Proteins to Functional Biomaterials. *Matter* **2020**, *2* (1), 50–77.

https://doi.org/10.1016/j.matt.2019.11.011.

(47) Gunnoo, S. B.; Madder, A. Chemical Protein Modification through Cysteine. *ChemBioChem* **2016**, *17* (7), 529–553. https://doi.org/10.1002/cbic.201500667.

(48) Asiimwe, N.; Al Mazid, M. F.; Murale, D. P.; Kim, Y. K.; Lee, J. S. Recent Advances in Protein Modifications Techniques for the Targeting N-Terminal Cysteine. *Pept. Sci.* **2021**, No. June. https://doi.org/10.1002/pep2.24235.

(49) Potts, K. T.; Baum, J. S. The Chemistry of Cyclopropenones. *Chem. Rev.* **1974**, *74* (2), 189–213. https://doi.org/10.1021/cr60288a003.

(50) Lovely, L. Squeeze Play. *MSW Manag.* **2017**, *27* (3), 28–31.

(51) Reber, K. P.; Gilbert, I. W.; Strassfeld, D. A.; Sorensen, E. J. Synthesis of (+)-Lineariifolianone and Related Cyclopropenone-Containing Sesquiterpenoids. *J. Org. Chem.* **2019**, *84* (9), 5524–5534. https://doi.org/10.1021/acs.joc.9b00478.

(52) Breslow, R.; Haynie, R.; Mirra, J. The Synthesis of Diphenylcyclopropenone [3]. *J. Am. Chem. Soc.* **1959**, *81* (1), 247–248. https://doi.org/10.1021/ja01510a060.

(53) W. Mercer, S. Winn, H. W. Twinning in Crystals of Human Skeletal Muscle D-Glyceraldehyde-3-Phosphate Dehydrogenase. **1976**, *575* (2), 3215.

(54) Jenkins, J. L.; Tanner, J. J. High-Resolution Structure of Human D-Glyceraldehyde-3-Phosphate Dehydrogenase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (3), 290–301. https://doi.org/10.1107/S0907444905042289.

(55) Guido, R.; Balliano, T.; Andricopulo, A.; Oliva, G. Kinetic and Crystallographic Studies on Glyceraldehyde-3-Phosphate Dehydrogenase from Trypanosoma Cruzi in Complex with Iodoacetate. *Lett. Drug Des. Discov.* **2009**, *6* (3), 210–214. https://doi.org/10.2174/157018009787847774.

(56) Park, J. B.; Park, H.; Son, J.; Ha, S. J.; Cho, H. S. Structural Study of Monomethyl Fumarate-Bound Human Gapdh. *Mol. Cells* **2019**, *42* (8), 597–603. https://doi.org/10.14348/MOLCELLS.2019.0114.

(57) Chen, Z.; Li, Y.; Sun, X.; Yuan, Q. Improvement of Expression Level of Polysaccharide Lyases with New Tag GAPDH in E. Coli. *J. Biotechnol.* **2016**, *236*, 159–165. https://doi.org/10.1016/j.jbiotec.2016.08.016.

(58) Schmidt, T. G. M.; Skerra, A. The Strep-Tag System for One-Step Purification and High-Affinity Detection or Capturing of Proteins. *Nat. Protoc.* **2007**, *2* (6), 1528–1535. https://doi.org/10.1038/nprot.2007.209.

(59) Istrate, A.; Navo, C. D.; Sousa, B. B.; Marques, M. C.; Deery, M. J.; Bond, A. D.; Corzana, F.; Jiménez-Osés, G.; Bernardes, G. J. L. Selective N-Terminal Cysteine Protein

Modification with Cyclopropenones. *ChemRxiv* **2020**, 1–23.

(60) Ii, P. S.; Strep-tactin, T. Strep-Tactin ® Purification - Short Protocol - Strep-Tactin ® Purification - Short Protocol -. **2020**, No. February, 1–4.

(61) Healthcare, G. E. PD-10 Desalting Columns.

(62) Pak; , William L.; Grossfield, Joseph; Arnold, K. S. © 1970 Nature Publishing Group. *Nat. Publ. Gr.* **1970**, *228*, 726–734.

(63) Carpenter, D. K. Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics (Berne, Bruce J.; Pecora, Robert). *J. Chem. Educ.* **1977**, *54* (10), A430. https://doi.org/10.1021/ed054pa430.1.

(64) Borgstahl, G. E. O. How to Use Dynamic Light Scattering to Improve the Likelihood of Growing Macromolecular Crystals. *Methods Mol. Biol.* **2007**, *363*, 109–129. https://doi.org/10.1385/1-59745-209-2:109.

(65) Bergsdorf, C.; Wright, S. K. *A Guide to Run Affinity Screens Using Differential Scanning Fluorimetry and Surface Plasmon Resonance Assays*, 1st ed.; Elsevier Inc., 2018; Vol. 610. https://doi.org/10.1016/bs.mie.2018.09.015.

(66) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat. Protoc.* **2007**, *2* (9), 2212–2221. https://doi.org/10.1038/nprot.2007.321.

(67) Naik, S.; Zhang, N.; Gao, P.; Fisher, M. T. *On the Design of Broad Based Screening Assays to Identify Potential Pharmacological Chaperones of Protein Misfolding Diseases*; 2013; Vol. 999. https://doi.org/10.2174/15680266112129990072.

(68) Miles, A. J.; Wallace, B. A. Circular Dichroism Spectroscopy for Protein Characterization: Biopharmaceutical Applications. *Biophys. Charact. Proteins Dev. Biopharm.* **2015**, 109–137. https://doi.org/10.1016/B978-0-444-59573-7.00006-3.

(69) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2007**, *1* (6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.

(70) Bulletin, T. Technical Bulletin Maxiren ® Technical Bulletin Maxiren ®. No. 2 mM, 1–4.

(71) White, M. R.; Khan, M. M.; Deredge, D.; Ross, C. R.; Quintyn, R.; Zucconi, B. E.; Wysocki, V. H.; Wintrode, P. L.; Wilson, G. M.; Garcin, E. D. A Dimer Interface Mutation in Glyceraldehyde-3-Phosphate Dehydrogenase Regulates Its Binding to AU-Rich RNA. *J. Biol. Chem.* **2015**, *290* (3), 1770–1785. https://doi.org/10.1074/jbc.M114.618165.

(72) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data Processing and Analysis with the AutoPROC Toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 293–302.

https://doi.org/10.1107/S0907444911007773.

(73) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674. https://doi.org/10.1107/S0021889807021206.

(74) Potterton, L.; McNicholas, S.; Krissinel, E.; Gruber, J.; Cowtan, K.; Emsley, P.; Murshudov, G. N.; Cohen, S.; Perrakis, A.; Noble, M. Developments in the CCP4 Molecular-Graphics Project. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12 I), 2288–2294. https://doi.org/10.1107/S0907444904023716.

(75) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. Biological Crystallography A Graphical User Interface to the CCP4 Program Suite. *Acta Cryst* **2003**, *59*, 1131–1137.

(76) Cowtan, K. The Buccaneer Software for Automated Model Building. 1. Tracing Protein Chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (9), 1002–1011. https://doi.org/10.1107/S0907444906022116.

(77) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12 I), 2126–2132. https://doi.org/10.1107/S0907444904019158.

(78) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53* (3), 240–255. https://doi.org/10.1107/S0907444996012255.

(79) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221. https://doi.org/10.1107/S0907444909052925.

(80) DeLano, W. L. Pymol: An Open-Source Molecular Graphics Tool. *CCP4 Newsl. protein Crystallogr.* **2002**, *40*, 82–92.

(81) Diederichs, K.; Karplus, P. A. Improved R-Factors for Diffraction Data Analysis in Macromolecular Crystallography. *Nat. Struct. Biol.* **1997**, *4* (4), 269–275. https://doi.org/10.1038/nsb0497-269.

(82) Matthews, B. W. Solvent Content of Protein Crystals. *J. Mol. Biol.* **1968**, *33* (2), 491–497. https://doi.org/10.1016/0022-2836(68)90205-2.

(83) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **2019**, *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

(84) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

(85) Park, J. B.; Park, H.; Son, J.; Ha, S. J.; Cho, H. S. Structural Study of Monomethyl Fumarate-Bound Human GAPDH. *Mol. Cells* **2019**, *42* (8), 597–603. https://doi.org/10.14348/molcells.2019.0114.

(86) Dierks, K.; Meyer, A.; Einspahr, H.; Betzel, C. Dynamic Light Scattering in Protein Crystallization Droplets: Adaptations for Analysis and Optimization of Crystallization Processes. *Cryst. Growth Des.* **2008**, *8* (5), 1628–1634. https://doi.org/10.1021/cg701067r.

(87) Torres-Bugeau, C. M.; Ávila, C. L.; Raisman-Vozari, R.; Papy-Garcia, D.; Itri, R.; Barbosa, L. R. S.; Cortez, L. M.; Sim, V. L.; Chehín, R. N. Characterization of Heparin-Induced Glyceraldehyde-3-Phosphate Dehydrogenase Early Amyloid-like Oligomers and Their Implication in α-Synuclein Aggregation. *J. Biol. Chem.* **2012**, *287* (4), 2398–2409. https://doi.org/10.1074/jbc.M111.303503.

(88) Provencher, S. W. A Constrained Regularization Method for Inverting Data Represented by Linear Algebraic or Integral Equations. *Comput. Phys. Commun.* **1982**, *27* (3), 213–227. https://doi.org/10.1016/0010-4655(82)90173-4.

(89) Micsonai, A.; Wien, F.; Bulyáki, É.; Kun, J.; Moussong, É.; Lee, Y. H.; Goto, Y.; Réfrégiers, M.; Kardos, J. BeStSel: A Web Server for Accurate Protein Secondary Structure Prediction and Fold Recognition from the Circular Dichroism Spectra. *Nucleic Acids Res.* **2018**, *46* (W1), W315–W322. https://doi.org/10.1093/nar/gky497.

(90) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. The Proteomics Protocols Handbook. *Proteomics Protoc. Handb.* **2005**, 571–608. https://doi.org/10.1385/1592598900.

(91) Dimensions, M. RUBIC Buffer Screen For Stable , Happy Proteins – From Purification All the Way through to Characterization by NMR , SAXS or Crystallography . **2013**, 1–7.

(92) Research, H. Izit Crystal Dye User Guide. **1995**, *85* (2206), 2.

(93) Rossman, M. G.; Liljas, A.; Brändén, C. I.; Banaszak, L. J. 2 Evolutionary and Structural Relationships among Dehydrogenases. *Enzymes* **1975**, *11* (C), 61–102. https://doi.org/10.1016/S1874-6047(08)60210-3.

(94) Butterfield, D. A.; Hardas, S. S.; Lange, M. L. B. Oxidatively Modified Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and Alzheimer's Disease: Many Pathways to Neurodegeneration. *J. Alzheimer's Dis.* **2010**, *20* (2), 369–393. https://doi.org/10.3233/JAD-2010-1375.

Supplementary Information

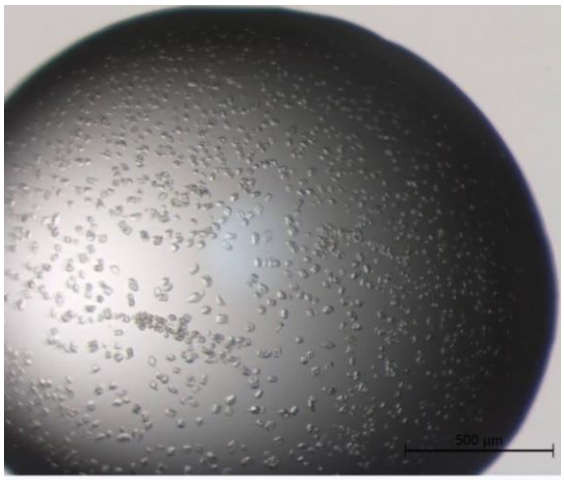

Figure S.1. First crystals obtained, later found to be salt crystals. The drop consisted of 0.1 µL of GAPDH protein (10 mg/mL) and 0.1 µL of screen solution: 20% (w/v) PEG 3350, 0.02 M $ZnCl_2$.

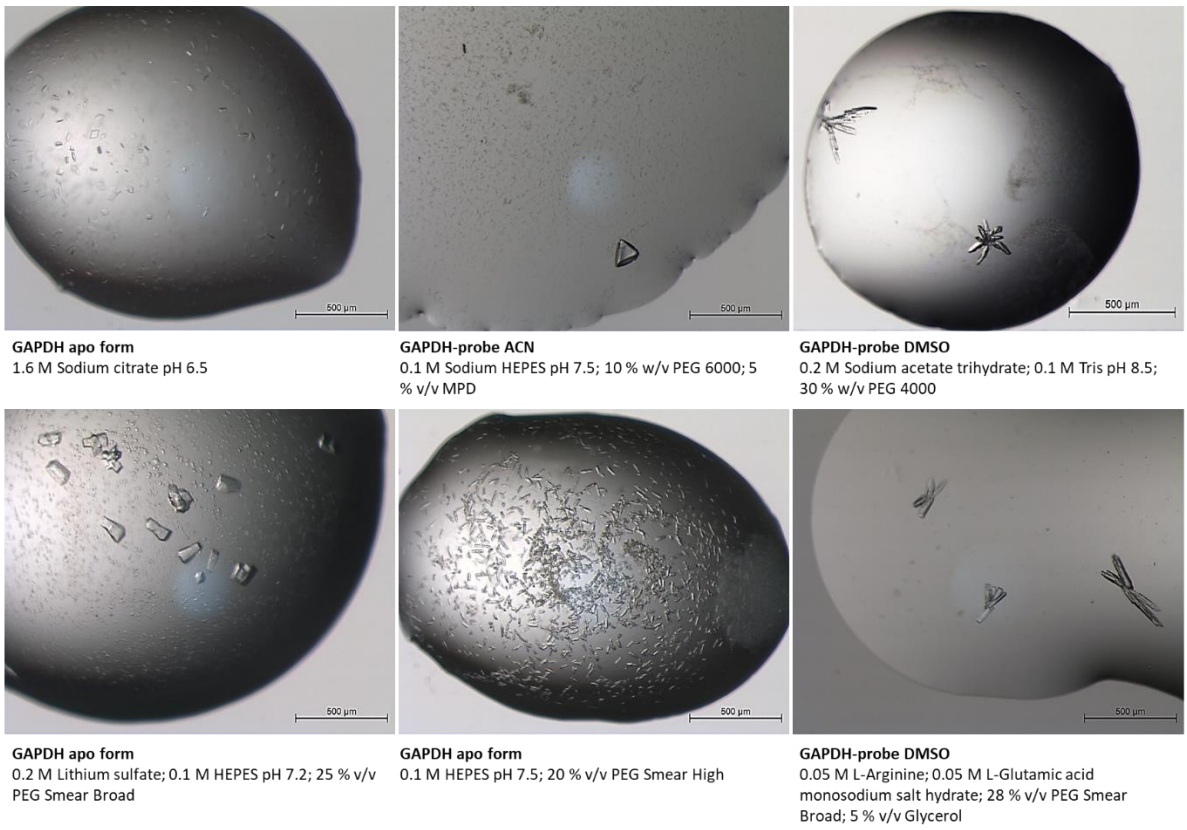

Figure S.2. Crystals obtained from the first Screen 1 and 2 and BCS screen by carrying out the experiments on a crystallization robot. Crystallization trials with hGAPDH (10 mg/mL) alone – apo form – or pre-incubated with CPO-probe in ACN and in DMSO.

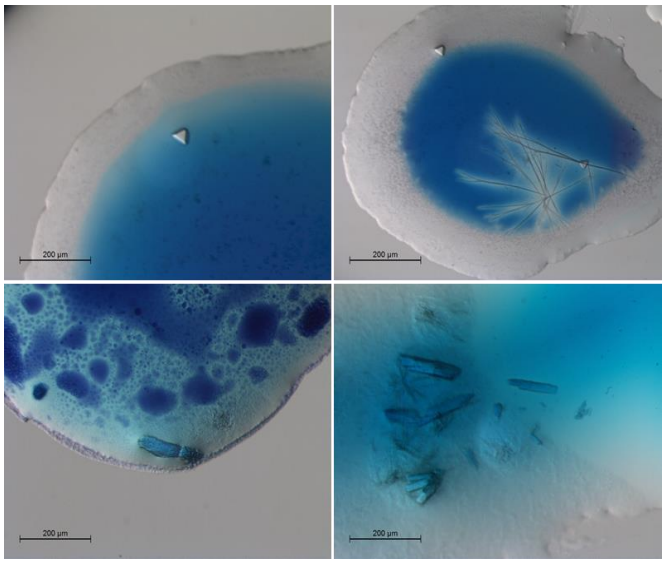

Figure S.3. Izit crystal dye assay performed in the crystals obtain by S1 and 2 and BCS, shown previously. The protein crystals are colored blue and salt crystals are colorless.

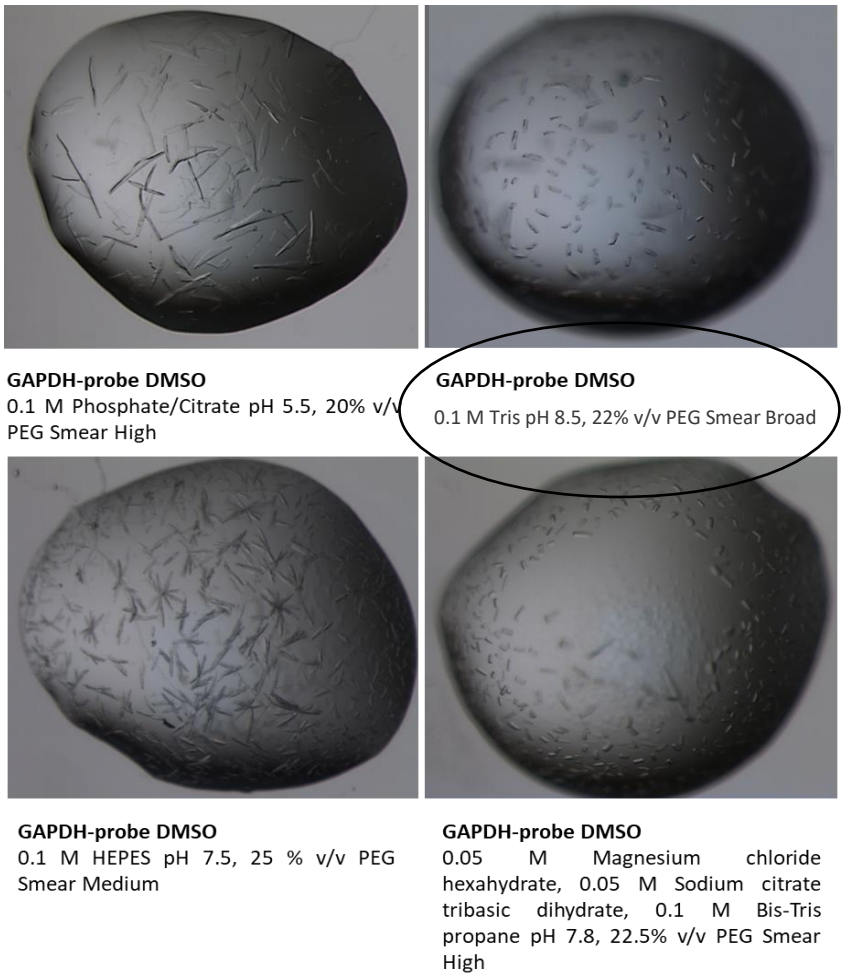

Figure S.4. Crystals obtain from the first Screen 1 and 2 and BCS screen by carrying out the experiments on a crystallization robot. Crystallization trials with hGAPDH in complex with CPO-probe in DMSO. The most promising result is marked in figure by a circle.