



DIPLOMARBEIT

Titel der Diplomarbeit

Protein Folding Mechanism

Analysing the Folding Mechanism of myPGK Utilizing Simulations and Experiments

Verfasser

Wolfgang Schlattl

angestrebter akademischer Grad

Magister der Pharmazie (Mag.pharm.)

Wien, 2014

Studienkennzahl:

A 449

Studienrichtung:

Diplomstudium Pharmazie

Betreuer:

Univ.-Prof. Mag. Dr. Gerhard F. Ecker

This work is dedicated to my parents

Acknowledgement

Hereby, I would like to express my sincere gratitude to all people who supported me and gave me the opportunity to realize this elaborate topic.

Very special thanks go to my supervisor Prof. Gerhard F. Ecker for his support, very helpful suggestions and the critical review of this work. Furthermore, I want to extend my particular appreciation to Prof. Gottfried Köhler from the Department of Structural and Computational Biology for his effort to enable all experiments, the suggestion to explore the properties of the modified yeast phosphoglycerate kinase and his helpful interpretation of the results. I also want to thank Prof. Peter Wolschann from the Institute for Theoretical Chemistry for his encouragement in order to perform the simulations, his advices and his support to get access to the Vienna Scientific Cluster 1.

Moreover, I appreciate the help from all people of the Pharmacoinformatics Research Group. Special thanks are due to DI(FH) Dr. Daniela Digles, Mag.pharm. Dr. Barbara Zdrazil and Amir Seddik, MSc. who explained to me various aspects of molecular dynamics simulations. I would like to give special thanks to Chonticha Suwattanasophon, MSc. from the Theoretical Biochemistry Group for the introduction to GROMACS and the support to handle the simulations. Furthermore, I want to thank her for the help to perform simulations on the VSC 1. I also want to thank all people of the Biomolecular Optical Spectroscopy Group. I am especially grateful due to the extraordinary teamwork with Dipl. Ing. Arthur Sedivy, his introductions to most of the instruments which were used for the experiments, his effort to gain the vector of the modified yeast phosphoglycerate kinase from the Forschungszentrum Jülich and his mentoring. Furthermore, I would like to thank Mag. Martin Gerald Puchinger for his helpful suggestion during all experiments. Thanks go to Mag. Erwin Gaubitzer for his support concerning several issues with different software packages and the introduction to circular dichroism. Special thanks go to Dr. Ulrich Salzer who introduced me to many molecular biological methods and techniques. I also want to thank Claudia Schreiner from the Department of Structural Biology of Cytoskeleton for her help to manage the molecular biological experiments. Special thanks go to Dr. Katarzyna Cieřlik-Boczula who helped me to acquire the necessary experience in order to perform

reproducible DSC records. Besides, I want to thank Dr. Markus Stöhr for his support to get Desmond running on the VSC-1. Furthermore, I am very grateful for the fruitful cooperation with the Forschungszentrum Jülich. I would like to thank especially Dr. Tobias Rosenkranz for his work and the molecular cloning of the modified yeast phosphoglycerate kinase vector. I also want to thank Prof. Dr. Jörg Fitter who agreed to the cooperation and gave us access to the vector. Special thanks go to Dr. Tamas Vidoczy from the Institute of Structural Chemistry of the Hungarian Academy of Sciences who made it possible to perform the fluorescence time life measurements.

At last, I want to thank my parents D.I. Manfred Schlattl and Maria Schlattl for their support, trust and faith in me all over the years. Additionally, I want to thank my brother Gerhard for all his advices and all my friends who helped me to come this far.

Of course I want to thank my better half Verena for correcting this work and for her moral support.

Table of content

Acknowledgement.....	II
Table of content.....	IV
1 Introduction	7
1.1.1 History	7
1.1.2 Goal	12
1.1.3 Model enzyme – yeast phosphoglycerate kinase	13
2 Materials.....	15
2.1.1 Instruments and components	15
2.1.2 Cell lines.....	16
2.1.3 Kits and enzymes	16
2.1.4 Vector	17
2.1.5 Chemicals	17
2.1.6 Media, buffer and solutions.....	20
3 Methods.....	23
3.1 Experimental methods.....	23
3.1.1 Molecular biological methods.....	23
3.1.1.1 Transformation	23
3.1.1.2 Amplification	24
3.1.1.3 Gel electrophoresis	24
3.1.1.4 Glycerol stock	26
3.1.1.5 Cell disruption – dynamic high pressure homogenizer (DHPH)	26
3.1.1.6 Ammonium sulphate precipitation	27
3.1.1.7 Immobilized metal chelate affinity chromatography (IMAC)	28
3.1.1.8 Fast protein liquid chromatography (FPLC) – ÄKTA	29
3.1.2 Biochemical methods	30
3.1.2.1 Enzyme activity.....	30
3.1.2.2 Labelling with extrinsic fluorophores	32
3.1.3 Optical methods.....	32
3.1.3.1 Absorption spectroscopy	32

3.1.3.2	Circular dichroism spectroscopy (CD)	35
3.1.3.3	Fluorescence spectroscopy	39
3.1.3.4	Fluorescence lifetime measurements	44
3.1.3.5	Förster resonance energy transfer - FRET	45
3.1.4	Calorimetric method	50
3.1.4.1	Differential Scanning Calorimetry (DSC)	50
3.2	Simulation methods	54
3.2.5	Homology Modelling.....	54
3.2.6	Molecular Dynamics Simulations (MD).....	55
3.2.6.1	Leap frog integrator	58
3.2.6.2	Periodic boundary condition.....	58
3.2.6.3	Steep descent minimization	58
3.2.6.4	Linear Constraints Solver (LINCS)	58
3.2.6.5	Berendsen thermostat.....	59
3.2.6.6	Velocity-rescaling thermostat	59
3.2.6.7	Parrinello-Rahman barostat	59
3.2.7	Determination of the orientation factor - κ^2	60
4	Results.....	61
4.1	Experiment.....	61
4.1.1	Amplification	61
4.1.2	Expression.....	64
4.1.3	Purification.....	64
4.1.4	Labelling of myPGK.....	68
4.1.4.1	Packing a DEAE sepharose column	68
4.1.4.2	Evaluation of the DEAE sepharose column	68
4.1.4.3	Labelling procedure for the double labelled myPGK.....	70
4.1.4.4	Labelling procedure for acceptor/donor – only myPGK	73
4.1.5	Structural and functional validation (CD, enzyme activity, DSC)	75
4.1.5.1	Circular dichroism (CD)	75
4.1.5.2	Enzyme activity	77
4.1.5.3	Differential Scanning Calorimetry (DSC)	78

4.1.6	Fluorescence and absorption measurements	81
4.1.7	Quantum yield and spectral overlap	84
4.1.8	Efficiency of energy transfer – EET	85
4.1.8.1	EET from fluorescence intensities	85
4.1.8.2	EET from fluorescence lifetime	86
4.2	Simulation	88
4.2.9	Homology Modelling	88
4.2.10	Molecular dynamic simulation	89
4.2.10.1	GROMACS (wtPGK, myPGK)	90
4.2.10.2	Desmond (myPGKdl)	94
4.2.11	Orientation factor – κ^2	97
4.3	Analysis	99
4.3.12	Förster radius – R_0	99
4.3.13	Temperature dependent distance distribution - r	99
5	Discussion	101
6	Conclusion and Outlook	102
7	Abbreviations & Mathematical Terms	103
8	References	107
9	Declaration	112
10	Appendix	113
10.1.1	myPGK sequence	113
10.1.2	Modeller – Homology modelling alignment	113
10.1.3	GROMACS – setup	114
11	Abstract	119
12	German abstract - Zusammenfassung	120
13	Curriculum Vitae	121

1 Introduction

1.1.1 History

In the last centuries several scientific milestones have been accomplished in the field of cellular research. This success ranges from the first description of the cell by Robert Hooke in the year 1665 (Bayrhuber et al. 1996) until the latest discovery of the 'machinery regulating vesicle traffic' by J. E. Rothman, R. W. Schekman and Th. C. Südhof who were awarded the Nobel Price in Physiology or Medicine 2013 (Nobel Media AB 2013).

Nevertheless, there are still scientific questions which need to be answered. One of them is the folding problem of proteins. Basically, proteins are polymers of 20 different amino acids which are linked by peptide bonds. Furthermore, the sequence of this polypeptide defines the primary structure. Each primary structure is unique and essential to form the correct three-dimensional structure and its related function. This connection was first described by Christian B. Anfinsen (Nobel Prize 1972) in the 1960s (Lajtha & Banik 2007). He and his team worked with ribonuclease A which consists of 124 amino acids and contains 4 disulfide bonds. They found out that the protein acquires its native conformation after renaturation and formed the 4 correct disulfide bonds out of 105 possibilities. The conclusion was that protein folding is a reversible and spontaneous reaction. Based on this knowledge, the following questions arose: How does the mechanism of protein folding work and how is this information encoded in the protein sequence? Anfinsen argued that folding of an unfolded de novo protein is thermodynamically driven by the force of a free energy gradient until it reaches the lowest energetic conformation ($\Delta G = 0$). This gradient is supported by the hydrophobic effect and non covalent interactions like hydrogen bonds and electrostatic interactions. The only counteracting force is based on the conformational entropy loss due to reduced flexibility (Müller-Esterl, 2004). Besides, Cyrus Levinthal published his work in the year 1968 which became known as the Levinthal paradox. It states that the time for protein folding could be estimated, if a random search of all possible backbone conformations is done (Buchner & Kiefhaber 2005). Consequently, if a protein consists

of 100 amino acids and each of them could obtain only two energetically preferred conformations, there would be 2^{100} or 10^{30} possibilities in total. The transition from one conformation to another takes 10^{-11} s which means that the complete search would take 100 billion years (Müller-Esterl, 2004). However, proteins fold in a time range of microseconds to minutes. The conclusion of this paradox was that there must be intermediates and pathways to reach the energetic favourable conformation in time (Buchner & Kiefhaber 2005). As a result, several protein folding models were established to describe the folding pathways. Four of the classic models are depicted in Figure 1.1.

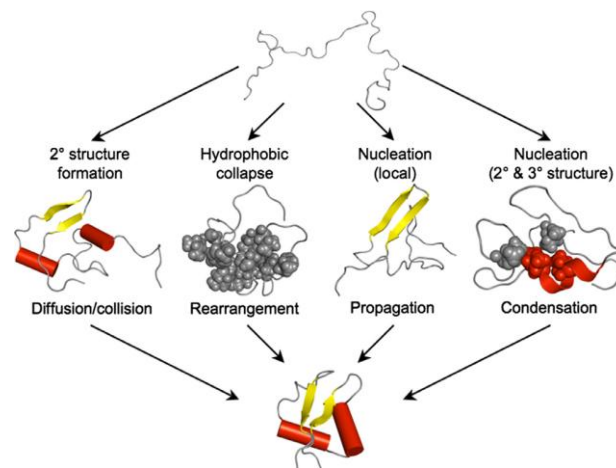


Figure 1.1: Protein folding models, from left to right: framework model, hydrophobic collapse model, nucleation-propagation model, nucleation-condensation model (Nickson & Clarke 2010)

In the framework model the protein folds according to a hierarchical scheme. Therefore, it is assumed that the unfolded protein (random coil) forms secondary structures at first and finally accomplishes the native state by diffusion, collisions and adherence. In the hydrophobic collapse model the hydrophobic side chains of the protein rapidly form a so called 'molten-globule' intermediate before the tertiary structure is generated. The basic idea behind the nucleation-propagation model is the formation of small amounts of secondary structures because of local interactions. Consequently, the protein folding prolongs nearby these elements until the structure is completed. The last model is the nucleation-condensation model. The folding mechanism of this model starts with a metastable nucleus which is unable to perform further folding until a critical number of long range interactions are established. From that moment on, the tertiary structure gets

formed so fast that the nucleus can not be completed in the transition state (Nickson & Clarke 2010). Besides these classic models a jigsaw puzzle model was published in 1985. It proposes that the protein does not follow a specific pathway to reach its final conformation. On the contrary, this model suggests that each protein molecule is able to find its native structure on a separate way. This hypothesis led to a new perspective of protein folding which is in accordance with the Gibbs free energy landscape. The energy landscape is a mathematical function with n degrees of freedom which can be used to describe molecular properties of a system. In case of a protein each energy function $F(x) = F(x_1, \dots, x_n)$ represent the free energy of a specific conformation. In addition, the variables x_i can be described as individual dihedral angles of the backbone and the entropic part of the free energy $F(x)$ results from the sum of all solvent configurations without the chain conformations. Consequently, the native state of the protein is defined by the set of individual conformations (x_1, \dots, x_n) with the lowest free energy value. Usually, an energy landscape is a multi-dimensional surface. However, they are commonly transformed to a 3D model for a better visualization. In this model the vertical axis equals the free energy of the system and the horizontal axes correspond to the conformational space of the protein backbone. These properties led to the conclusion that the shape of the energy surface is comparable to a funnel as illustrated in Figure 1.2 (Lajtha & Banik 2007).

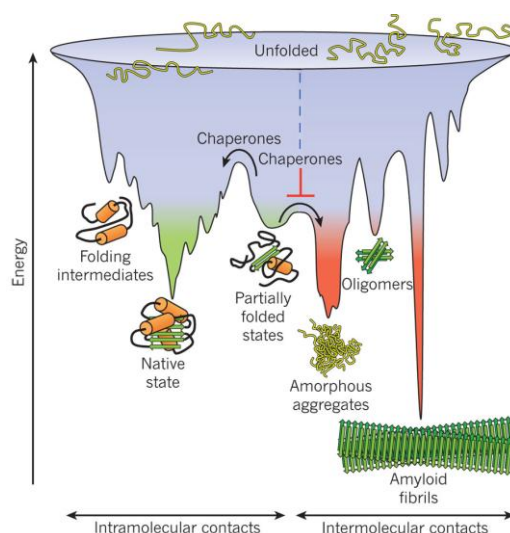


Figure 1.2: Protein folding funnel (Hartl et al. 2011)

The surface of this folding funnel can be relatively rough because of several local minima. These local minima can also be traps for further folding and the protein remains in a misfolded structure or starts to aggregate. However, the backbone of evolutionary optimised proteins fold relatively fast into the native low energy conformation. This point in the energy landscape is also described as 'minimal frustrated' and represents the global minimum. In contrast, at the top of the funnel is a plateau of high free energy which is characterised by a variety of unfolded protein forms. Consequently, several pathways can lead the unfolded protein downwards to the native structure by reducing its free energy and the number of possible conformations. This concept gives also a plausible answer to the question: Why is the structure of a protein relatively robust to mutations? It is because this perturbation has little impact on the shape of the energy surface and the location of the global minimum. This conclusion supports also the principle of homology modelling which is used to predict protein structures (Lajtha & Banik 2007).

The folding funnel delivers an excellent explanation to describe protein folding but until now there is not a practical, accurate and universal method which could be used to determine the tertiary structure based on the protein sequence. In order to solve this problem it would be necessary to know all positions of every protein atom and surrounding atoms for every time step during protein folding. Therefore, several theoretical and experimental methods have been used to describe this process in the last decades. On the theoretical side molecular dynamics simulations at the *ab initio* level would be accurate enough to generate a model based on physical laws. Although this method is part of an active research field it requires a vast amount of computational resources and provides minimal output. Consequently, this technique is not yet feasible. On the experimental side equilibrium measurements and kinetic measurements are utilized to gather information about the folding process. Equilibrium measurements are used to find possible folding intermediates and to determine rate constants from molecular fluctuations (Lajtha & Banik 2007). The principle of kinetic measurements is to monitor the system after a rapid change of an extrinsic property like pH, temperature, pressure or salt concentration until it reaches the equilibrium again. The detection methods for such techniques are based on fluorescence, phosphorescence, circular dichroism, infrared spectroscopy, nuclear magnetic resonance and mass spectroscopy.

Generally, the time of protein folding depends on the protein and ranges from microseconds to minutes. Nowadays, in-silico methods are only practical to analyse folding mechanisms up to 1ms. Besides, it is rather difficult to gain an experimental detection method with an adequate time resolution and without a dead time (Lajtha & Banik 2007). As a result, single methods are not sufficient yet to analyse the whole folding process in detail. Therefore, simulations are used to create models of the protein folding process which are validated by experiments or vice versa.

Most experiments are performed in vitro to acquire physical properties. In the cell protein folding is even more complex because it starts during the expression and the final structure can only form afterwards. Small proteins fold spontaneously and larger ones are supported by peptidyl-prolyl-cis/trans isomerases, disulfide isomerases or chaperones in order to form correctly. Nevertheless, the results of in vivo measurements are relevant especially to understand misfolding, aggregation and related diseases. Such protein associated diseases are the result of different pathological mechanisms. If too few functional proteins are in the cell it could lead to cystic fibrosis and early-onset emphysema. Insufficient protein elimination is often the reason for protein aggregation. The resulting β -amyloides and plaques cause neurodegenerative diseases like Alzheimer's, Parkinson's, Huntington's disease and spongiform encephalopathies like BSE and scrapie. Protein deposition in the Langerhans's islands is known to cause diabetes II and other diseases are associated with protein misfolding like the Creutzfeldt-Jacob syndrome, AL amyloidosis and amyotrophic lateral sclerosis (Müller-Esterl, 2004) (Lajtha & Banik 2007).

1.1.2 Goal

This work represents participation in a project called 'Refolding of Biological Macromolecules Initiated by Ultrafast cooling'. The goal of the project is to analyse the thermal refolding process of denatured proteins by smFRET. Therefore, the aim of this thesis is to create reference data by determining the thermal unfolding profile of C97S Q135C S290C phosphoglycerate kinase (myPGK) from *Saccharomyces cerevisiae* between 15 °C – 75 °C by utilizing Förster resonance energy transfer (FRET). The resulting experimental data will be corrected by simulations and compared to the literature. Finally, reference data will be generated of the temperature dependent distance distribution between the two domains of the protein. Further structural properties will be acquired by circular dichroism (CD) and differential scanning calorimetry (DSC). The general workflow of this thesis is depicted in Figure 1.3.

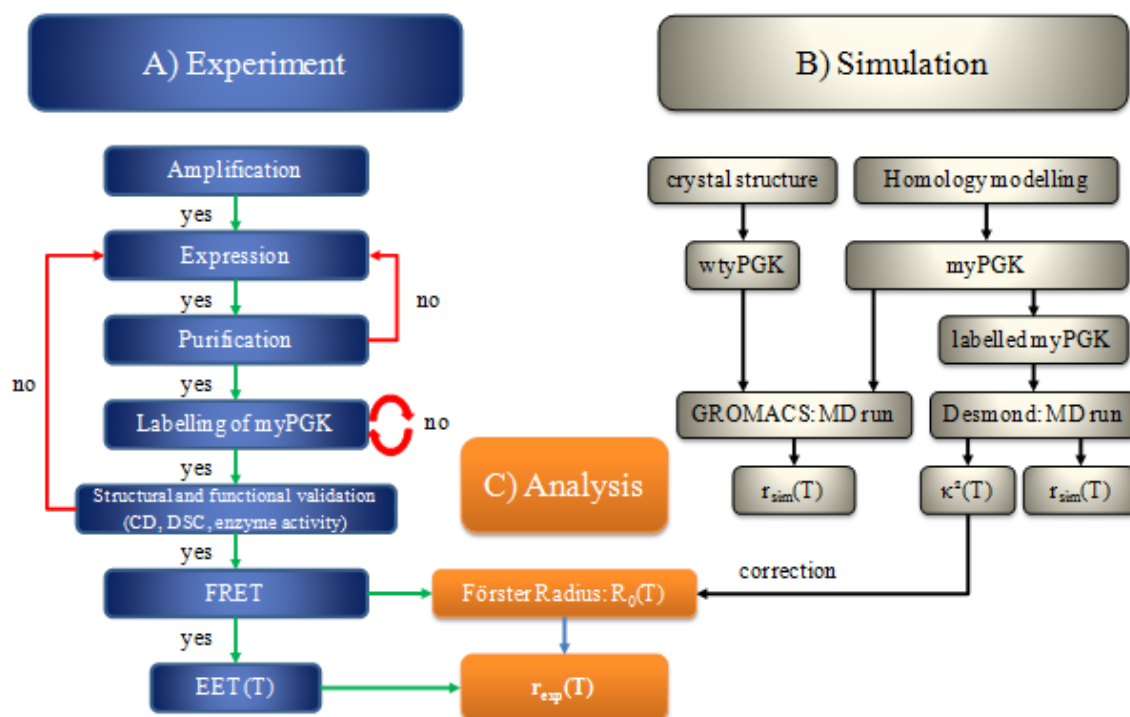


Figure 1.3: Workflow of the thesis, CD: circular dichroism, DSC: differential scanning calorimetry, FRET: Förster resonance energy transfer, EET: efficiency of energy transfer, wtyPGK: wild type yeast phosphoglycerate kinase, myPGK: modified yeast phosphoglycerate kinase, $r_{sim}(T)$: simulated interdomain distance, κ^2 : orientation factor, $r_{exp}(T)$: experimental determined interdomain distance

As illustrated in the workflow, the experiment starts with the amplification of the vector and continues with the expression and purification. In the next step, the myPGK becomes labelled with Alexa488 and Alexa647 in position 155 and position 310. Consequently, the structure and function of the protein gets validated by circular dichroism (CD), differential scanning calorimetry (DSC) and an enzyme activity assay. Afterwards, spectroscopic properties can be determined by measuring the fluorescence, absorption and reference index of the sample.

Besides, molecular dynamic simulations are performed with GROMACS of the wild type form (wtPGK) and myPGK to analyse the influence of the mutations to the interdomain distance. Hence, the structure of wtPGK is obtained from the Protein Data Bank (PDB) and the structure of myPGK is acquired by homology modelling. In addition, the dynamics of labelled myPGK is simulated with Desmond to get information about the distance between the dyes and to calculate a realistic orientation factor κ^2 .

Finally, the measured and simulated spectral properties are used according to the Förster formula to determine a realistic temperature dependent distance distribution between cysteine 155 of the N-terminal domain and cysteine 310 of the C-terminal domain of myPGK. The outcome will be discussed in comparison to the literature.

1.1.3 Model enzyme – yeast phosphoglycerate kinase

The wildtype form of yeast phosphoglycerate kinase is a monomeric protein of 415 amino acids (44 500 Da) which forms two domains of equal size. Generally, it is built up by repeating $\alpha\beta$ -motifs. Each domain consists of six parallel β -sheets that are surrounded by α -helices. Therefore, 25 % of the protein is part of β -sheets and 42 % of helices. Helix V and G β represent the connection between the domains which give the protein its characteristic movement of a hinge bending enzyme. Furthermore, the C-terminal chain is attached to the N-terminal domain. The main difference of the mutant is that it has 19 additional amino acids on the C- and N-terminal side which results in 454 amino acids (48 600 Da). In addition, a His-tag sequence is located at the beginning of the N-terminus to simplify the purification procedure after the expression. The last characteristic attributes are the C97S Q135C S290C mutations which are

introduced because the cysteines act as labelling position for the fluorescent dyes (Figure 1.4) (Yon-Kahn & Hervé 2010).

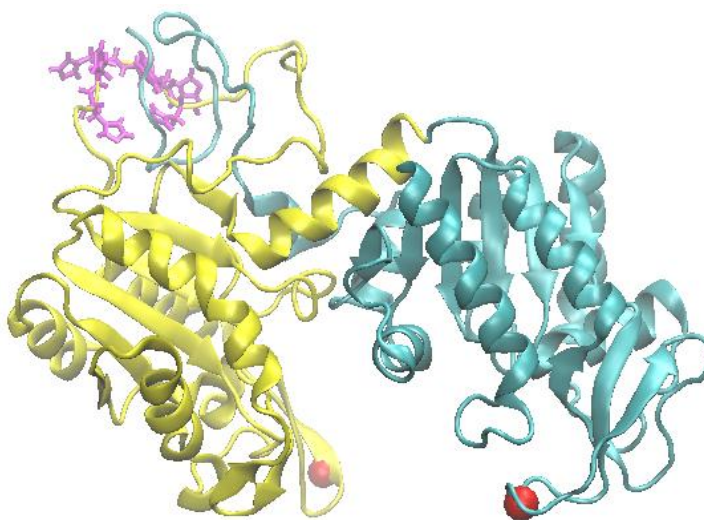
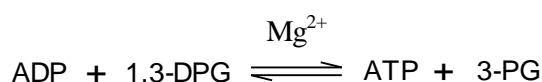


Figure 1.4: modified yeast phosphoglycerate kinase, N-terminal domain (yellow), C-terminal domain (cyan), labelling positions 155 and 310 (red), HIS-taq sequence (magenta) (created with Maestro 2012)

Generally, phosphoglycerate kinase is a kinase of the carboxylate type ($R-COO^-$) and it is the first enzyme of the glycolytic pathway which produces ATP. Therefore, it catalyses the reversible transfer of phosphate from 1,3-diphosphoglycerate (1,3-DPG) to ADP in the presence of Mg^{2+} and releases 3-phosphoglycerate (3-PG) and ATP (Yon-Kahn & Hervé 2010). This reaction is summarized below:



Several studies of the mechanism showed that the binding site of Mg-ATP or Mg-ADP is at the C-terminal domain and it is proposed that 3-phosphoglycerate binds to Arg38 and Arg168 of the N-terminal domain. The binding of 3-PG also induces a hinge bending motion and reduces the interdomain distance from 15 Å to 5 Å and its carboxy group forms a connection to the C-terminal domain via a water molecule. The domain movement seems to be the key for the catalytic circle. Unfortunately, there is still no proof of this theory because until now it was not possible to isolate the ternary complex of the substrates (Yon-Kahn & Hervé 2010).

2 Materials

2.1.1 Instruments and components

Äkta: Äkta purifier, Amersham pharmacia biotech, P-900, UV-900, pH/C-900, Box-900, P-960, Inv-907, PU-908, Frac-950 (System 2)

Electrophoresis: BioRad, MiniSub[®] Cell GT

BioRad, Mini Protean 3

BioRad, Power Pac HC 250V; 3,0A; 300W

Centrifugation concentrator: Vivaspin 20, Sartorius Stedim Biotech, 10.000 MWCO, prod#: VS 2001/2002

Vivaspin 500, Sartorius Stedim Biotech, 10.000 MWCO, prod#: VS0102

Centrifuges: Heraeus, Multifuge 3 L-R

Eppendorf, Minispin Plus 5415D

Sorvall RC 5c Plus, Rotors: SLC 4000, SS-34, GS-3

Circular Dichroism: CD - Instrument: Applied Photophysics π^* 180

Lamp: Xenon lamp, MA1343

Cuvette: Hellma, Suprasil[®] 110-QS 1.00 mm, 284

Desalting column: Disposable PD 10 Desalting Columns, GE Healthcare, 17-0851-01

Dialysis membrane: Roth, ZelluTrans Dialysiermembranen 3,5, MWCO: 4.000 – 6.000

DSC: MicroCal[™], VP-DSC MicroCalorimeter

MicroCal[™], ThermoVac

Fluorescence lifetime spectrometer:

Fluorescence lifetime spectrometer, Edinburgh Instruments, FL 900

Picosecond Pulsed Diode Laser, Edinburgh Instruments, EPL-470

Picosecond Pulsed Diode Laser, Edinburgh Instruments, EPL-635

Fluorimeter: Perkin Elmer, Luminescence Spectrometer, LS 50B

Fridge: Liebherr Premium KGT 3946 24D /001

New Brunswick, Ultra Low Temperature Freezer, U725 Innova

Homogenizer: Avestin, Emulsiflex C5

Incubator: Incucell MMM

Ion Exchange Column: GE Healthcare Column K9/30 (Code.No. 19-0871-01)

Microwave: micromaxx[®] MM 41580

Nanodrop: Thermo Scientific, 2000C, UV - Vis spectrophotometer

Ni-NTA column: Amersam Biosciences, HisTrap FF crude, 5 ml

Pump: Pharmacia LKB- Pump P-1

Refractometer: Arthago NAR- 1T 1210

Shakers: New Brunswick Scientific, Innova 4400 incubator shaker

Sartorius Stedim Biotech, Certomat[®] BS-1

UV/VIS – Spectrophotometer: Hitachi, U-3300 Spectrophotometer, INULA

Vortexer: IKA[®], MS2 Minishaker

2.1.2 Cell lines

E. Coli BL21 DE3 cells: Invitrogen, One Shot[®] BL21(DE3) Chemically Competent *E. coli*

E. Coli DH5 α cells: Invitrogen, MAX Efficiency[®] DH5 α TM Competent Cells

2.1.3 Kits and enzymes

BamH I: Fermentas, Fastdigest BamHI, # FD0054, 800 μ l

GAPDH: Glyceraldehyde 3 phosphate Dehydrogenase, from rabbit muscle, Sigma Life Science, CAS 9001-50-7, G2267-500UN, 109 U/mg protein

Miniprep Kit: GeneJetTM Plasmid Miniprep Kit, # K0502, # K0503, Fermentas

Nde I: Fermentas, Fastdigest Nde I, # FD0584, 300 μ l

wtyPGK: 3- Phosphoglyceric phosphokinase, (*S. cerevisiae*), Sigma Life Science, CAS: 9001-83-6; 3,9 mg protein/ml (WC): 1616 units/mg protein

2.1.4 Vector

pET 15b + PGK C97S Q135C S290C (= pET 15b + myPGK) vector:

The engineered vector was created by Rosenkranz T. (2011a) and obtained from Prof. Dr. Jörg Fitter of the Jülich Forschungszentrum. This vector consists of the gene modified yeast phosphoglycerate kinase and the pET-15b plasmid (Figure 2.1). This plasmid also includes important qualities like an ampicillin resistance gene, restriction sites for BamH I and Nde I enzymes and a His-tag sequence for the protein.

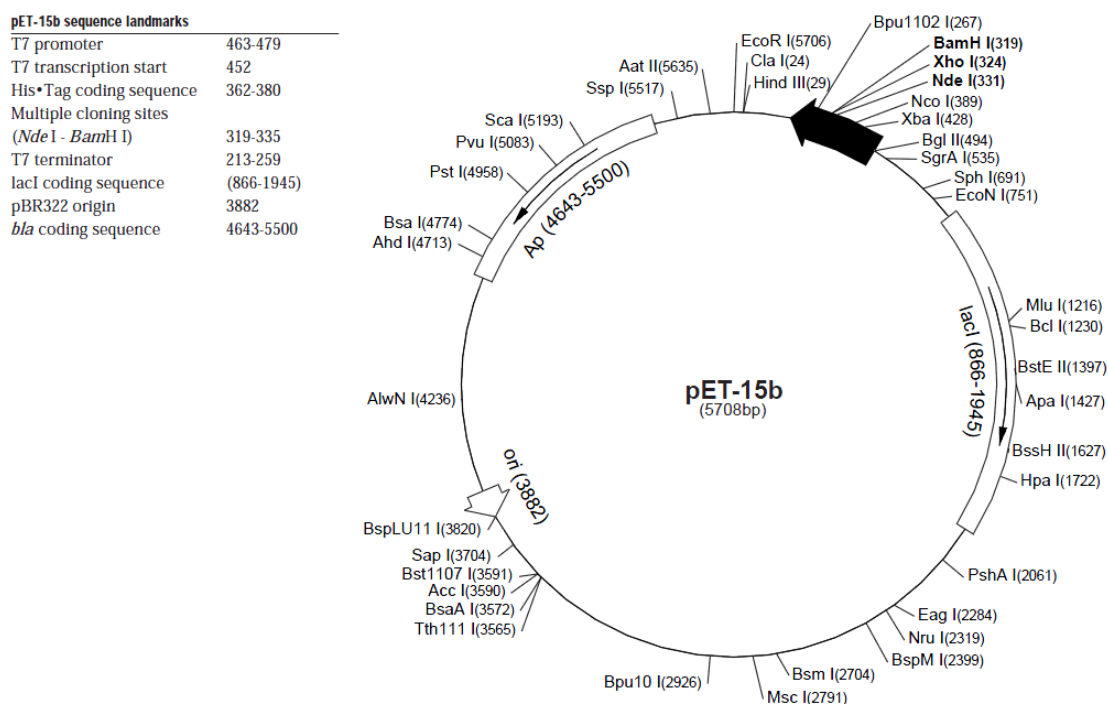


Figure 2.1: Genetic map of the pET-15b expression vector (Merck Millipore 2013)

2.1.5 Chemicals

Acetone: Acetone, Sigma Aldrich, CAS: 67-64-1, P#: 00570-2,5L

Acrylamide: Rotiphorese[®] Gel 30 (37,5:1), Roth, UN#: 3426

Agar: Select agar, Sigma Aldrich, A5054-1KG, CAS: 9002-18-0

Agarose: Agarose, For Routine Use, Sigma A 9539-500G, CAS 9012-36-6

Alexa488: Alexa Fluor[®] 448 C₅ Maleimide, Invitrogen, Cat#: A-10254

Alexa647: Alexa Fluor[®] 647 C₂ Maleimide, Invitrogen, Cat#: A-20347

Ammonium persulfate:	Ammoniumperoxodisulfat, ROTH, CAS: 7727-54-0, Art#: 9592.3
Ammonium sulphate:	Ammonium sulfate, Sigma Aldrich, CAS: 7783-20-2, P#: A4418-1KG
Bromphenol blue:	(BPB) Bromphenol Blue, plus one, Code#: 17-1329-01
BSA, purified:	100x concentrate, New England Biolabs Inc, Cat#: B9001S
Carbenicillin:	Carbenicillin disodium salt, AppliChem - Biochemica A1491,0010, CAS: 4800-94-6
Chloroform:	Chloroform, Sigma Aldrich, CAS: 67-66-3, P#: 372978-1L
Coomassie brilliant Blue:	Brilliant Blue R, Sigma Aldrich, CAS: 6104-59-2, P#: B7920-10G
Cysteine:	L-Cysteine, Sigma Aldrich, CAS: 52-90-4, P#: 168149-25G
Decon 90:	Zinsser Analytic, Konzentrat, 200ml
DNA Marker:	Gene Ruler™ 1 kb DNA Ladder, Ready for Use, Fermentas, 0,5 µg/µl, 50µg #SM0311
DTT:	DL-Dithiothreitol, Sigma Aldrich, CAS: 3483-12-3, P#: 43819-25G
EDTA:	Ethylenediaminetetraacetic acid tetrasodium salt dihydrate, Sigma Aldrich, Fluka BioChemika, CAS: 10378-23-1, prod#: 03695-250G
Ethanol:	Ethanol, Sigma Aldrich, CAS: 64-17-5, P#: 245119
Fluorescein:	Fluorescein sodium salt, Sigma Aldrich, CAS: 518-47-8, prod#: F6377
Glacial acetic acid:	Acetic acid, Sigma Aldrich, CAS: 64-19-7, P#: 695092-2,5L
Glycerol:	Glycerin, ROTH, ROTIPURAN®, ≥99,5 %, p.a., water free, CAS: 56-81-5
Glutathione:	L- Glutathione reduced, Sigma Aldrich, CAS 70-18-8, prod#: G4251-1G

Imidazole:	Imidazole, Sigma Aldrich, CAS: 288-32-4, P#: 56750-1KG
Ion exchange material:	GE Healthcare DEAE Sepharose Fast Flow (Code.No. 17-0709-10) 25 ml
IPTG:	Isopropyl β -D-1-thiogalactopyranoside, peqlab, Cat#: 37- 2030
Isopropyl alcohol:	2-Propanol, Sigma Aldrich, CAS: 67-63-0, prod#: 33539-1L
KH ₂ PO ₄ :	Potassium phosphate monobasic, Sigma Aldrich, CAS: 7778-77-0, P#: 60218-1KG
LB Broth:	LB- Broth, Sigma-Aldrich, L3022-1KG (10 g/l tryptone, 5 g/l NaCl, 5 g/l Yeast extract)
Methanol:	Methanol, Sigma Aldrich, CAS: 67-56-1, P#: 322415-1L
β -Mercaptoethanol:	2-Mercaptoethanol, Sigma Aldrich, CAS: 60-24-2, SKU: M6250
MgATP:	Adenosine 5'-triphosphate magnesium salt, Sigma Aldrich, CAS: 74804-12-9, prod#: A9187-100mg
MgSO ₄ :	Magnesium sulfate heptahydrate, Sigma Aldrich, CAS: 10034-99-8, P#: 63138-1KG
MOPS:	3-(N-morpholino)propanesulfonic acid, Sigma Aldrich, CAS: 1132-61-2, P#: M1254-250G
NaCl:	Sodium chloride, Sigma Aldrich, CAS: 7647-14-5, P#: 71376
β -NADH:	Beta Nicotinamide adenine dinucleotide sodium salt from <i>Saccharomyces cerevisiae</i> , Sigma Life Science, CAS: 20111-18-6, prod#: N0632-1G
NEBuffer4:	10x concentrate, New England Biolabs Inc, Cat#: B7004S
Nitrogen (liquid):	Messer Austria, Stickstoff 5.0, UN 1066
3PG:	DL Glyceraldehyde 3 phosphate solution, Sigma Aldrich, CAS: 591-59-3, prod#: G5251-25MG

PMSF:	Phenylmethanesulfonyl fluoride, Sigma Aldrich, CAS: 329-98-6, prod#: P7626-1G
Protein Marker:	PageRuler™ Plus Prestained Protein ladder, Thermo Scientific, P#: 22621, 25 µl
SDS:	Sodium dodecyl sulfate, Sigma Aldrich, CAS: 151-21-3, prod#: L3771-100G
Silica gel:	LUDOX® HS-30 colloidal silica, Sigma Aldrich, 30 %, prod#: 420824-1L
SYBR:	SYBR® Safe DNA gel stain, 10.000 x concentrate in DMSO, Invitrogen, Cat#: S33102
TCEP:	Tris(2- carboxyethyl) phosphinehydrochloride, Aldrich Chemistry, CAS 51805-45-9, prod#: C4706-10G
TEMED:	N,N,N',N'-Tetramethylethylenediamine, Roth, CAS: 110-18-9
Tripotassium phosphate:	Potassium phosphate tribasic, Sigma Aldrich, CAS 7778-532, prod#: P5629-500G
Tris:	Tris(hydroxymethyl)aminomethane, VWR International, CAS: 77-86-1, P#: 103156X

2.1.6 Media, buffer and solutions

LB medium:	20 g/l LB Broth (autoclaved)
Agar medium:	3 g LB Broth; 4,5 g Agar; 300 ml sterile water (autoclaved)
Carbenicillin solution:	0,1 g/ml Carbenicillin (sterilized by sterile filter)
Agar plates:	300 ml Agar medium were heated in the microwave until it was liquid. In addition, 300 µl carbenicillin solution was added in the cold but liquid agar medium and poured it into the plates.
Agarose gel electrophoresis:	TAE buffer: 48,4 g Tris, 11,4 ml glacial acetic acid, 3,7 g EDTA

Purification:

- Resuspension: 10 mM MOPS, 500 mM NaCl, 1 mM DTT, pH = 7,5
 Wash I: 10 mM MOPS, 10 mM Imidazole, pH = 7,5 (NaOH)
 Wash II: 10 mM MOPS, 20 mM Imidazole, pH = 7,5 (NaOH)
 Elution: 10 mM MOPS, 250 mM Imidazole, pH = 7,5 (NaOH)

SDS gel electrophoresis:

	running gel	stacking gel
	V [ml]	V [ml]
H ₂ O	4,0	3,4
30 % acrylamide mix (30 % + 0,8 %; acrylamide : N,N Methylenebisacrylamid)	3,3	0,83
1,5 M TRIS, pH = 8,8	2,5	0,63
10 % SDS	0,1	0,05
4 % APS	0,1	0,05
TEMED	0,004	0,004

- Sample buffer: 200 µl β Mercaptoethanol, 2 mg Bromophenol blue, 8 ml Lämmli buffer
 Lämmli buffer: 3 % TRIS HCl, 1 % SDS, 14,34 % Glycerol
 Staining solution: 0,25 % Coomassie Brilliant blue R250, 7,5 % glacial acetic acid, 50 % ethanol
 Destainer solution: 10 % glacial acetic acid, 30 % ethanol

Dialysis

- Folding Studies Buffer: 10 mM MOPS, 50 mM NaCl, 2 mM EDTA, pH = 7,4
 Labelling Buffer: 100 mM KH₂PO₄, 500 mM NaCl, 1 mM TCEP, pH = 7,5

Column packing:

- Packing buffer: 10 mM MOPS, 1 M NaCl
 Storage buffer: 10 mM MOPS, 20 % EtOH

Labelling:

- Phosphate buffer: 100 mM KH₂PO₄, 500 mM NaCl, pH = 7,5 (NaOH)
 MOPS buffer: 10 mM MOPS, pH = 7,4 (NaOH)
 Cleaning buffer: 10 mM MOPS, 1 M NaCl, pH = 7,4 (NaOH)
 NaOH: 1 M

3 Methods

3.1 Experimental methods

3.1.1 Molecular biological methods

All molecular biological methods were performed under sterile conditions!

3.1.1.1 Transformation

In nature genetic characteristics like antibiotic resistance or prototrophy can be transferred via vectors on a cellular level (Schlegel 1992). There are several ways to transmit genetic material from a donor to a recipient cell and some of them are conjugation, transduction and transformation. For the conjugation a so called mating bridge is formed between the cell-to-cell contact areas. This is necessary to establish the unilateral transfer of extra-chromosomal DNA through a protein channel. The horizontal transfer during the transduction depends on genetic vehicles like bacteriophages. The nucleotides within the phage are injected into the recipient cell after docking. The transformation includes the uptake of free DNA from the surrounding environment into competent bacterial cells. Afterwards, the DNA has to integrate into the chromosome to avoid degradation by nucleases. In practice, the gene of the desired protein is part of a plasmid and the transformation is induced for example by heat shock. Therefore, the membranes of the cells get permeable to free DNA without destroying too many of them (Dingermann et al. 2002).

Therefore, the transformation was used to amplify the pET 15b + myPGK vector stock with *E. Coli* DH5 α cells. Furthermore, it was the method of choice to prepare the expression of the myPGK in *E. Coli* BL21 DE3 cells. The transformation was induced by heat shock and executed according to the Stanford University (2003) protocol. First of all, the required competent cells were defrosted from the -80 °C freezer. In a 1,5 ml Eppendorf tube 50 μ l of the cells were mixed with 1 μ l (~50ng) of the solved vector and incubated for 10 minutes on ice. In the next step, the tube was heated for 45 s at 42 °C and then placed back on ice for 2 minutes. Afterwards 1ml of the LB medium (without antibiotics) was added to the tube and incubated for one hour at 37 °C and 300 rpm.

100 μ l of the cell suspension was spread over an agar plate and incubated upside down at 37 °C over night.

3.1.1.2 Amplification

The amplification can be described as the multiplication of DNA. Practically, there are different techniques to amplify nucleotides. For this work a cloning vector was utilized (Figure 2.1). Therefore, the gene of the desired protein was integrated in a small extra-chromosomal plasmid by ligation. The benefits of plasmids are their relative small size and with a sequence called ori (“origin of replication”) the replication does not depend on proliferation. Consequently, more copies of the same plasmid can be found in the cell in comparison to the large chromosome. The cloning vector also includes qualities to identify positive transformed cells like antibiotic resistance and many different restriction sites.

3.1.1.3 Gel electrophoresis

This molecular biological tool is an excellent analytical or preparative technique. In this work it was used to verify the amplification of the pET 15b + myPGK vector and to control the protein purification. The basic principle of electrophoresis is the migration of charged molecules in the electric field (Müller-Esterl, 2004). Practically, if a vertical gel electrophoresis is used, the gel stands between two glass plates in the anode buffer and is in contact with the cathode buffer on top. The sample gets applied into the pockets of the gel. The separation of different charged molecules starts after the connection to direct current of 120 – 250 V (Figure 3.1).

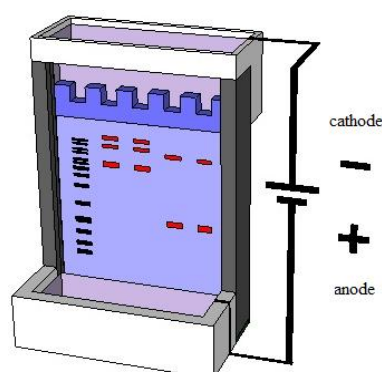


Figure 3.1: Vertical SDS PAGE – The marker (black bands) and sample (red bands) get applied to the pockets of the stacking gel (dark blue). Utilizing direct current separates a mixture of different charges molecules in the running gel (light blue).

This separation is the result of different migration velocities within the running gel. Depending on the molecule, the velocity increases proportional with the charge/mass ratio and decelerates with the size and shape. Consequently, DNA fragments can be separated according to their size because the structure is relatively rigid and the charge/mass ratio is constant. Generally nucleotides are pulled toward the anode considering the fact that the phosphate backbone is charge negatively at physiological pH (Müller-Esterl, 2004). The analysis of DNA over 25 kb starts with the fragmentation by restriction enzymes to enable the electrophoresis with a porous agarose gel (Schlegel, 1992). Thereby, smaller fragments migrate faster through the gel matrix than bigger ones. The resulting bands can be visualised after staining with the intercalating SYBR under UV light at 300 nm. Additional information can be obtained by creating restriction maps of digested and non digested DNA (Müller-Esterl, 2004).

In comparison to DNA or RNA the charge/mass ratios of proteins are erratic. To separate a complex protein mixture sodium dodecyl sulphate (SDS) (Figure 3.2) is used to circumvent this disadvantage.

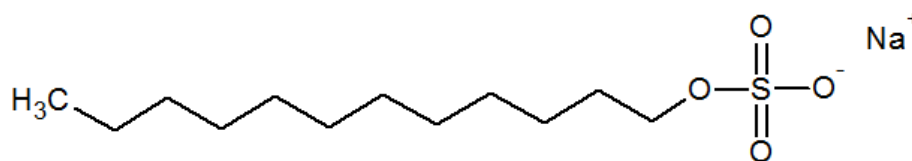


Figure 3.2: Sodium dodecyl sulphate (SDS), $\text{NaC}_{12}\text{H}_{25}\text{SO}_4$

SDS is an anionic detergent which denatures the proteins after binding with a great affinity. The stoichiometry is one SDS molecule to two amino acid residues. The attached SDS molecules overlay the charge of the protein and equalize the charge/mass ratios within a protein mixture. As a result the negative charged proteins migrate to the anode. The separation, during the SDS polyacrylamide gel electrophoresis (SDS PAGE), occurs according to the mass and is caused by the molecular sieve effect of the gel. The mesh size of this molecular sieve can be controlled by the concentration of the monomeric acrylamide and the cross-linker bisacrylamide. The separated proteins get selectively detected with Coomassie Brilliant Blue (Müller-Esterl 2004). Finally, the

relative mobility values can be calculated according to equation (3.1) and the masses are determined by comparison to the marker (Holtzhauer 2006).

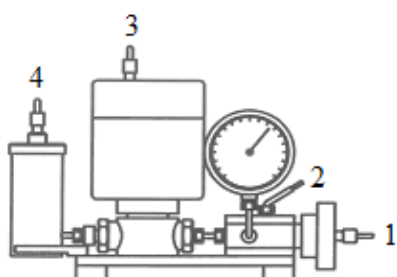
$$R_f = \frac{\text{distance}_{\text{macromolecule}}}{\text{distance}_{\text{trackingdye}}} \quad (3.1)$$

3.1.1.4 Glycerol stock

The glycerol stock is an easy way to store transformed cells at -80 °C. According to the protocol of the Hebrew University of Jerusalem (n.d.) 0,5 ml of 80 % glycerol get mixed with 0,5 ml of the transformed BL21 DE3 E. Coli cell culture (p. 23) in an Eppendorf tube. The tube was shock frozen with liquid nitrogen (-196 °C) and stored at -80 °C. New batches for expression were started by inoculating agar plates with the defrosted glycerol stock and they were incubated at 37 °C over night.

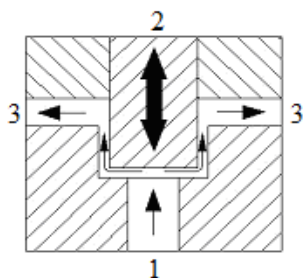
3.1.1.5 Cell disruption – dynamic high pressure homogenizer (DHPH)

In order to gain the protein after expression the cells have to be disrupted at first. Therefore, several techniques are available to perform this task and they can be subdivided into three classes: physical methods, chemical methods and mechanical methods (Sivasankar 2005). For this work the DHPH EmulsiFlex-C5 was used to mechanically disrupt the cells. The main components of this homogenizer are a gas driven high pressure pump and a pneumatically controlled, dynamic homogenizing valve. According to the special construction there is no need for gaskets in the entire product path (Avestin n.d.). The design of the used Emulsiflex-C5 is given in Figure 3.3.



*Figure 3.3: Elevation of EmulsiFlex-C5,
1) gas inlet for the high pressure pump,
2) sample outlet, 3) gas inlet for the
pneumatically controlled, dynamic
homogenizing valve, 4) sample inlet
(Avestin n.d, modified.)*

During the operation the cell suspension get pumped through the homogenizer with 15.000 psi (1034 bar). This pressure provides various stresses to the cell suspension and induces cell disruption at the dynamic homogenizing valve Figure 3.4.



*Figure 3.4: Schematically depiction of a dynamic homogenizing valve
1) sample inlet, 2) dynamic valve (plunger),
3) sample outlet (Avestin Inc n.d., modified)*

The main effect seems to be the interaction of the high velocity jet of the suspended cells with the stationary surface. This stress is described as dynamic pressure and depends on the jet velocity and the fluid density. Additional stresses are the normal and shear stress. The normal stress occurs during the fluid passes through the narrow channel and the shear stress happens because the pressure drops rapidly after the valve. The degree of cell disruption depends on several parameters like the nature of the cell, the product location within the cell, the used homogenizer, the pressure, the temperature and the number of passes through the homogenizer. Therefore, the sample was cycled 20 minutes through the Emulsiflex-C5 to be on the safe side. Heating of the cell suspension was avoided because the homogenisation was performed in a 4 °C cold room (Sivasankar 2005).

3.1.1.6 Ammonium sulphate precipitation

This is a crude but simple method to purify a protein after the cell disruption. The separation is the result of differences in solubility and is induced by increasing the ionic strength with an effective salt. The effectiveness of the salt follows the Hofmeister series. This series orders the kosmotropic effect of salts. It depends mostly on the anion which follow the order phosphate > sulphate > acetate > chloride. For the cation it is $\text{NH}_4^+ > \text{K}^+ > \text{Na}^+$. Ammonium sulphate is widely used because phosphates remain mostly as less effective H_2PO_4^- and HPO_4^{2-} and the solubility of potassium salts is poor.

Furthermore the salt is cheap and the density and temperature of the solution changes insignificantly after dissolving. The solubility of a protein in aqueous solution depends on the amount of hydrophilic and hydrophobic groups on its surface. The charged residues favour to interact with the polar solution. Nevertheless, there are also hydrophobic groups on the surface which are important for precipitation. If ammonium sulphate is added, the order of water molecules increases by solvation of the ions. Therefore, less water molecules are surrounding the protein and the hydrophobic areas on the surface get revealed. Consequently, these areas of the proteins interact with each other and induce precipitation by aggregation. This effect starts earlier for proteins which have more lipophilic groups on the surface. Therefore, the proteins get fractionated by increasing the salt concentration stepwise and each precipitate gets removed by centrifugation. Usually, the fractionation is performed at 4 °C in a buffer to avoid inactivation by proteases and precipitation because of a changing pH. At the end, the protein gets recovered by dissolving the pellet in an appropriate buffer and the salt is removed by dialysis (Roe 2001).

3.1.1.7 Immobilized metal chelate affinity chromatography (IMAC)

This method is used for a selective purification of proteins. The principle is based on the formation of a metal chelate complex. This complex is the result of a central transition metal ion which interacts non-covalently with ligands (Figure 3.5). Therefore, the stationary phase of the column consists of immobilised chelating ligands which are loaded with Ni^{2+} ions. These transition metal ions have additional binding sites for ligands like histidines of a protein. The bond to a single histidine is relatively weak and instable. Consequently, most proteins do not retain on the stationary phase because the histidines in most proteins are stochastically distributed. If, however, the protein has a His-tag sequence attached to the C- or N- terminus, which consists of several histidines in a row, the affinity to the Ni^{2+} increases exponentially. After the separation, the bound protein gets eluted by a high concentration of imidazole which binds competitive to the transition metal ion (Müller-Esterl 2004).

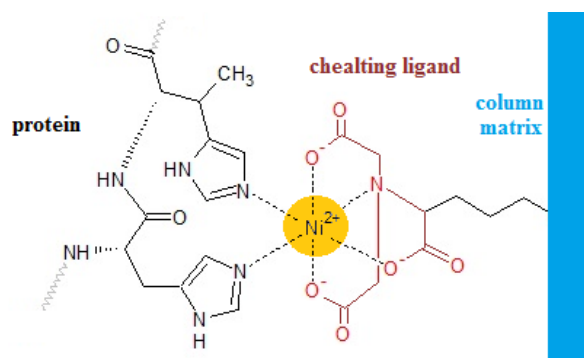


Figure 3.5: Protein – nickel - nitrilotriacetic acid (NTA) chelate complex (Müller-Esterl 2004, modified)

3.1.1.8 Fast protein liquid chromatography (FPLC) – ÄKTA

FPLC is based on the same principle like HPLC and was developed 1982 for the biocompatible high resolution separation of biopolymers (Cutler 2004). The main differences to HPLC are that higher protein loadings are separated at higher flow rates, with lower pressure (0,1 – 2 MPa) and in aqueous buffers. Furthermore, this method supports several chromatographic techniques to separate proteins, e.g. ion exchange, chromatofocusing, gel filtration, exclusion, affinity and reverse phase (Ghosal & Srivastava 2009).

For this work the FPLC was used in combination with a DEAE ion exchange (IEX) column. After the packing, it was integrated into the FPLC system as illustrated in Figure 3.6. At first, the column was evaluated by the detection of 0,2 ml of a 1 % acetone solution at 280 nm. Therefore, the system was flushed with double distilled water before the acetone sample was injected. Finally, the FPLC-IEX setup was used to separate unlabelled, single labelled und double labelled proteins during the labelling procedure. This was accomplished by equilibrating the system with MOPS buffer (solution A), injecting the labelled protein mixture via loop and by a stepwise increase of the Cleaning buffer concentration (solution B) (0 %, 4 %, 11,5 %, 18,5 %, 100 %, each 40 ml). Several parameters of the eluate were detected after the column like the UV at 260 nm, 280 nm and 647 nm, the conductivity and the pH. A detailed description of the column evaluation and the labelling procedure is on page 68.

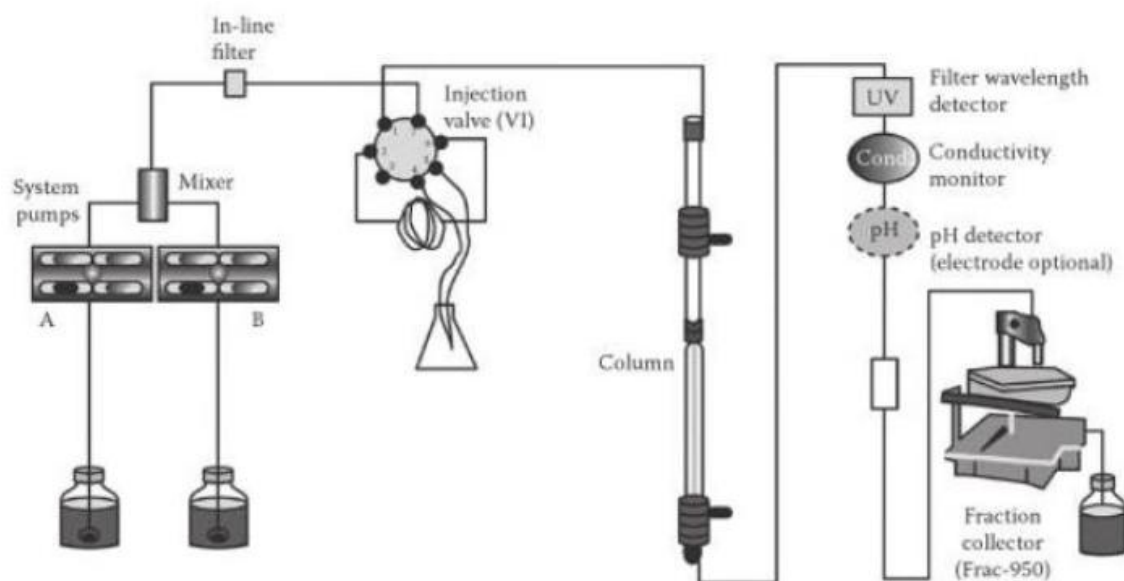


Figure 3.6: Schematic depiction of an FPLC system (ÄKTA) (Bisen & Sharma 2013)

3.1.2 Biochemical methods

3.1.2.1 Enzyme activity

The determination of the catalytic enzyme activity is a fast and selective way to confirm the native structure and function of a protein. In case of the modified yeast phosphoglycerate kinase (myPGK) the function was tested via a coupled optical test (Rosenkranz 2011a). Therefore, two reaction steps of the gluconeogenesis were combined to monitor the time dependent conversion of NADH to NAD^+ at 340 nm (Figure 3.7). The optical test was first described 1936 by the noble price winner Otto H. Warburg (1931) (Löffler et al. 2007).

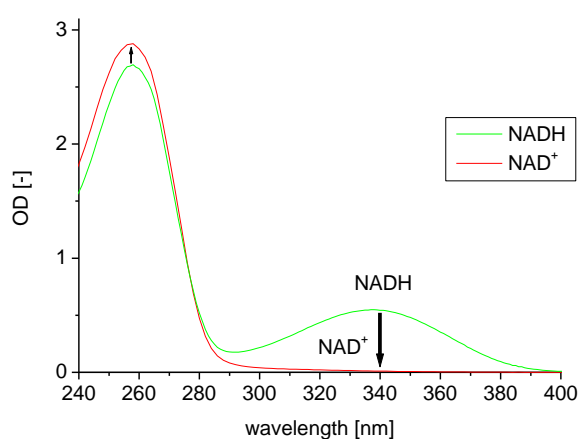
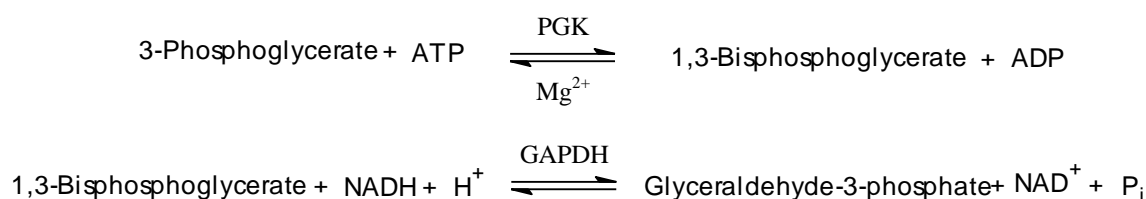


Figure 3.7: optical test – conversion of NADH to NAD^+

At the beginning, PGK catalyses the phosphorylation of 3-phosphoglycerate (3-PG) with ATP in the presence of magnesium ions to 1,3-bisphosphoglycerate. This product represents the substrate of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH). In the second reaction (indicator reaction) the 1,3-bisphosphoglycerate gets dephosphorylated and reduced by the cofactor NADH to glyceraldehyde-3-phosphate (Müller-Esterl 2004). Therefore, the concentration of NADH decreases over time as well as the absorption at 340nm. The reactions are summarised below.



For the experimental measurements all components of Table 3.1 were prepared except PGK. It was added at last to start the reactions before the absorption has been recorded for three minutes. The first three to five seconds got lost before the spectrometer started to measure.

Table 3.1: Composition of the enzyme activity assay

50 mM MgATP	100 µl
50 mM Cysteine	100 µl
250 mM 3-PG	20 µl
2 mM NADH	100 µl
GAPDH	11 U
100 mM MgSO ₄	50 µl
0,45 mM PGK	100 µl
Phosphate buffer	Up to 1000 µl

The decreasing absorption of NADH at 340 nm correlates directly to the concentration of converted 3-phosphoglycerate. The enzyme activity of such a kinetic can be described by the reaction velocity V . Generally, V is defined at optimal conditions by the micromoles of substrate which get converted within a minute. Therefore, the corresponding unit of the reaction velocity is U (= µmol min⁻¹). The International System of Units (SI) uses the unit katal (kat). It did not become generally accepted because 1 kat = 6 · 10⁷ U = 1 mol s⁻¹ (Richter 2003).

3.1.2.2 Labelling with extrinsic fluorophores

In most cases the protein of interest has not enough suitable intrinsic fluorophores like tryptophans or tyrosines for spectroscopic measurements. Hence, it is possible to attach extrinsic probes to the side chains of amino acids. Therefore, the PGK mutant with the desired labelling positions 155 and 310 was generated by site directed mutagenesis during the work of Dr. Rosenkranz (2011a). As a result, the maleimide groups of the Alexa dyes bind selectively by Michael addition to the mercapto groups of the cysteines (Figure 3.8).

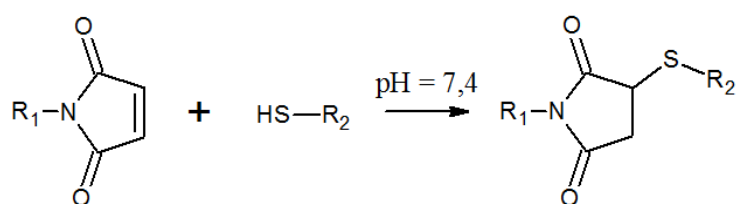


Figure 3.8: Labelling: Maleimide- Thiol reaction (Michael addition), R_1 : dyes (Alexa 488, Alexa 647), R_2 : cysteines of the protein (Hermanson 2008, modified)

The labelling procedure was performed in a 100 mM phosphate buffer at $\text{pH} = 7,5$ to avoid cross reactions with amines (Hermanson 2008). Furthermore, the buffer was degassed by nitrogen for 15 minutes to remove any interfering oxygen and a reductive environment was created by adding the thiol free TCEP to the solution (Rosenkranz 2011a). The labelling reaction was stopped with 3 mg L-glutathione after 12 hours.

This procedure was done to obtain double labelled myPGK (donor + acceptor), donor only myPGK and acceptor only myPGK. A detailed description of the labelling process is on page 68.

3.1.3 Optical methods

3.1.3.1 Absorption spectroscopy

The absorption of light by matter is a basic element of many spectroscopic techniques. As a result, multiple methods were established to observe different aspects of molecular structures. Some of them are listed in Table 3.2 (Galla 1988).

Table 3.2: Spectroscopic methods which utilize the absorption of electromagnetic radiation.

wavelength [nm]	spectral range	Method	effect on ...
1 – 10	X- ray	X- ray absorption	electron distribution
$10^2 - 10^3$	UV/VIS	UV/VIS absorption	
$10^3 - 10^5$	IR	IR/ Raman	vibration state
10^7	microwaves	EPR	electron spin
10^8	microwaves	microwave spectroscopy	molecule rotation
$>10^9$	radio waves	NMR	nuclear spin

(Galla 1988, modified)

For this work the spectral range of interest is the ultraviolet (UV) and visible (VIS) range. Generally, the Planck relation (3.2) postulates that light of a certain wavelength λ gets absorbed only if its energy E promotes an electron from the ground state to the excited one (Galla 1988).

$$E = h \cdot \nu = \frac{h \cdot c}{\lambda} \quad (3.2)$$

This relation of energy in J and the wavelength or frequency ν in s^{-1} is connected to the Planck constant $h = 6,63 \cdot 10^{-34}$ Js and the speed of light $c = 2,998 \cdot 10^8$ m \cdot s^{-1} , which are natural constants. The electron transitions are classified by the corresponding molecular orbital (Figure 3.9) (Galla 1988).

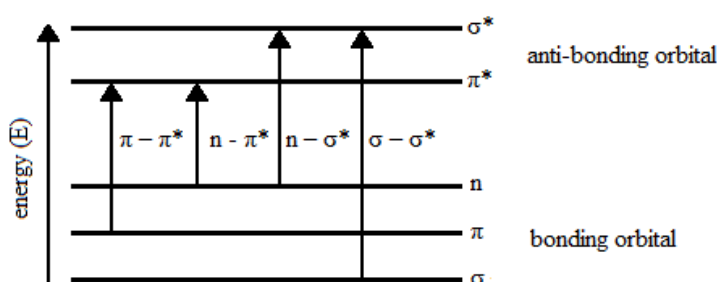


Figure 3.9: Energy levels of molecular orbitals and possible electronic transitions after absorption. (Galla 1988, modified)

In the ground state there are bonding σ -, π - and n - orbitals. Occupied bonding σ -orbitals form single bonds and occupied π -orbitals form multiple bonds within a molecule. Heteroatoms like N or O have also non-bonding orbitals which are occupied by lonesome electron pairs. From these orbitals of the ground state electrons can be transferred to empty anti-bonding orbitals. There are two types σ^* - and π^* - orbitals.

These energetic levels are the result of different molecular vibrations like stretching, bending and the rotation of covalent bonds. Important for bio molecules are the energetic lower $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions (Galla 1988). The absorption of light correlates to the exponential decrease in light intensity I and can be calculated according to the Lambert-Beer law:

$$OD = \log\left(\frac{I_0}{I}\right) = \varepsilon \cdot c \cdot d \quad (3.3)$$

where OD is the optical density at a defined light path d of 1 cm, c is the concentration in mol l^{-1} and ε in $\text{l mol}^{-1} \text{cm}^{-1}$ is the molar extinction coefficient of a substance. This exponential decay is measured by a spectrophotometer (Figure 3.10).

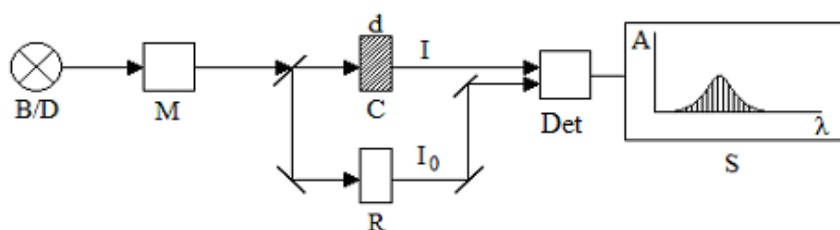


Figure 3.10: Assembly of a spectrophotometer. Light source B: bulb, D: deuterium lamp, M: monochromator, d: light path, I: intensity of monochromatic light from the sample cell C, I_0 : intensity of light from the reference cell R filled with buffer, Det: detector, S: spectrum

Measurements in the visible area (VIS) are made to determine the sample concentration or to observe the change in absorption of substrate coupled enzyme kinetic assays (p. 77). In the UV the temperature dependent unfolding within the proteins are measured by circular dichroism (p. 75) (Galla 1988). For this work all protein concentrations are measured and calculated according to the Warburg and Christian (WC) equation (3.4) (Holtzhauer 1997). This method has the advantage to measure the absorption of amino acids like tyrosine and tryptophan at 280 nm and correct the result for nucleic acid contamination at 260 nm (California Lutheran University 2012).

$$\beta_{WC} = 1,55 \cdot A_{280}^{1cm} - 0,76 \cdot A_{260}^{1cm} [\text{mg/ml}] \quad (3.4)$$

3.1.3.2 Circular dichroism spectroscopy (CD)

CD spectroscopy is in comparison to NMR spectroscopy and X-ray crystallography a relatively fast method to estimate the secondary structure of proteins in solution. Within a few hours it is possible to determine the influence of conformational changes because of mutations, unfolding or binding mechanisms. The principle of this method is defined by the difference of absorption between left-handed and right-handed circularly polarized light by chiral structures. This concept is based on the fact that light can be described as time dependent sinusoidal oscillating electric (and magnetic) fields which are isotropic distributed perpendicular to the direction of propagation. After the polarisation by filters the electric fields oscillate in a single plane. The linearly polarized light can be described as a linear combination of left and right handed circularly polarized light with equal amplitudes. If the amplitudes become unequal, because of different absorption, the resultant will oscillate elliptically (Figure 3.11) (Greenfield 2006).

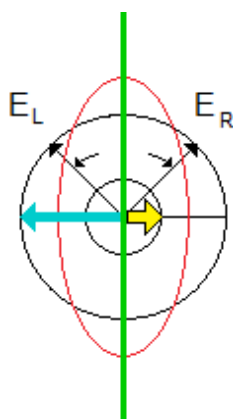


Figure 3.11: Linearly polarized light (green) is the resultant of left handed and right handed circularly polarized light with equal amplitudes (black arrows). If the lengths of both vectors are unequal (blue, yellow arrows), because of different absorption, the outcome will be ellipticity θ (red) (Galla 1988, modified).

Furthermore circularly polarized light requires a second linearly polarized light beam which is perpendicular to the first one and 90° out of phase (Figure 3.12-A). This is achieved by prisms or an electronic device using Pockel's effect¹. The resultant of the two linearly polarized beams is a left handed ($\Delta\phi = -90^\circ$) or a right handed ($\Delta\phi = +90^\circ$) circularly polarized light (Figure 3.12-B). With the view directly to the light source the

¹ Pockel's effect:

This ... electrically controllable ... effect (also called linear electro-optic effect) produces a change in the refractive index for light of certain polarizations which depends linearly on the strength of the applied (electric) field and it induces a change in the amount of birefringence present in an optical material like lithium niobate (LiNbO_3) (Drake 1996, p. 816).

vector of the resultant appears to rotate with a constant length in circles counter-clockwise or clockwise (Figure 3.12-C 0-3) (Greenfield 2006).

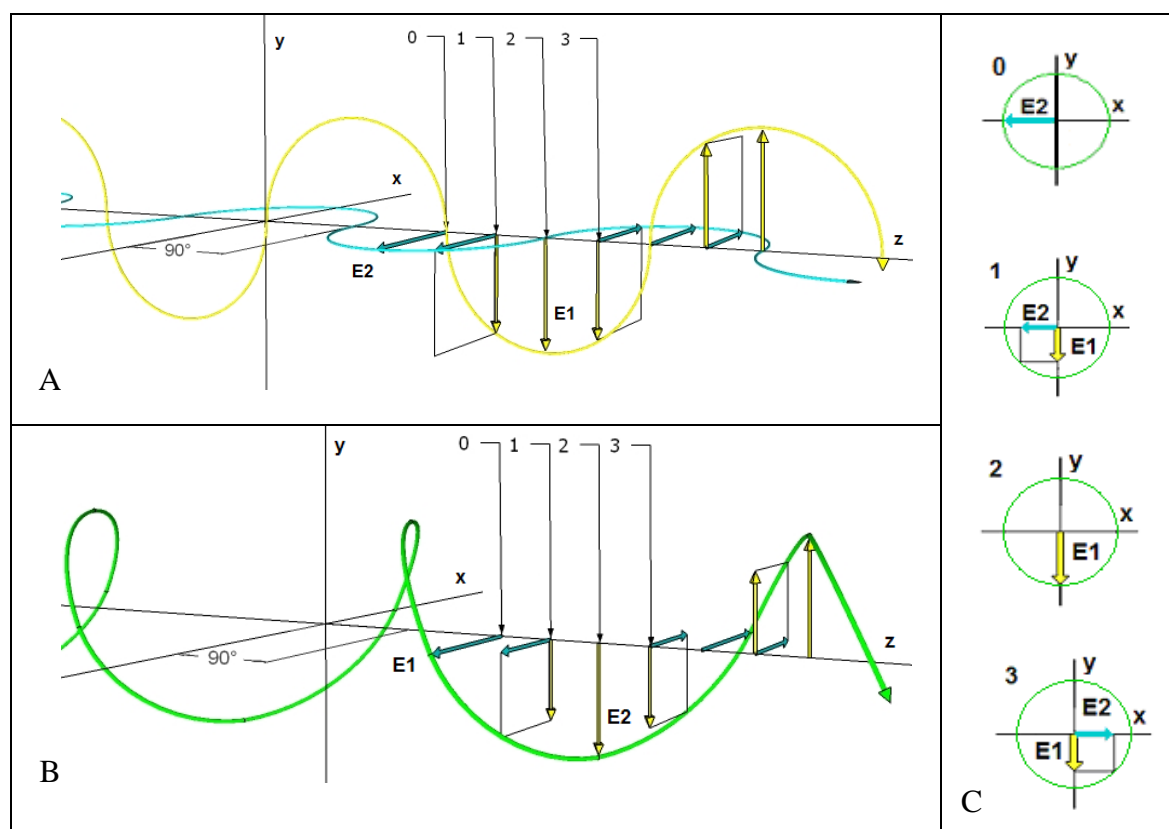


Figure 3.12: *Right circularly polarized light. Picture (A) shows the electronic vectors E1 and E2 of the two linearly polarized light beams which are perpendicular to each other and 90° ($\lambda/4$) out of phase in the direction of propagation. In picture (B) the right handed circularly polarized light is the resultant of the two polarized light waves. It rotates clockwise watching in the direction of the light source. The Pictures (C 0-3) display several snapshots of time and the resultant of the two vectors along the direction of propagation (Fasman 1996, modified).*

This circularly polarized light interacts with chiral structures like α -helices, β -sheets or with the backbone of a protein and gives characteristic spectra in the far (170 - 250 nm) and near (250 - 300 nm) UV. The absorption in the far UV depends mostly on the presence of amide bonds. Therefore, it is called amide region. Nevertheless, the aromatic side chains are also part of the signal in this region. In the near UV the signal originates mainly from the presence of aromatic side chains and a little from disulfide bonds. The CD spectrum in the far UV is the result of the absorption of all asymmetric structures and can be used to calculate or estimate the structural composition of a

protein. In most cases the information of the near UV, also called the fingerprint region, is inconclusive. It is only useful to compare spectra (Buchner & Kiefhaber 2005).

In practice, CD spectra are recorded by a CD spectrometer and a typical assembly is shown in Figure 3.13. At the beginning, a light source (e.g. Xenon arc lamp) sends a polychromatic light beam to the monochromator to get light of the desired wavelength. In the next, step the isotropic light passes through a polarizer to generate linearly polarized light. A CD modulator manipulates this light by inducing a phase shift of 90° like a quarter-wave plate ($\lambda/4$). These plates generate a right-handed circularly polarized light if the polarisation plane enters 45° inclined from the optical axis. Then the quarter-wave plate gets turned to an angle of 315° ($= -45^\circ$). The induced phase shift switches from $\lambda/4$ to $-\lambda/4$ and generates a left-handed circularly polarized light. Nowadays, the mechanical movement of a plate is obsolete because of modern CD modulators. They use the effect of applied electric fields to change the molecular orientation of the crystal. As a result, the switch between the two phase shifts and the direction of the circularly polarized light depends directly on the frequency of the alternating current. These two light beams are sent alternatively through a thin quartz cuvette which contains the chiral sample. Hence, the resulting signals get amplified by a photomultiplier and the outcome is represented in difference of absorption ΔA or ellipticity θ in degrees against the wavelength in nm (Galla 1988).

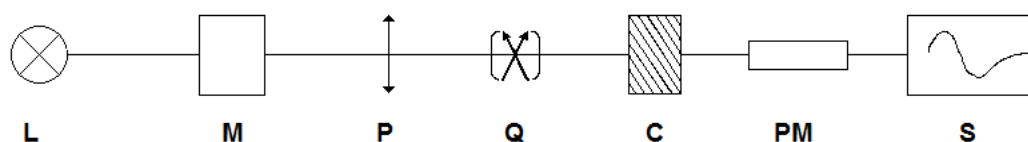


Figure 3.13: Assembly of a CD spectrometer, L: light source, M: monochromator, P: polarizer, Q: CD modulator, C: cuvette, PM: photomultiplier, S: spectrum (Galla 1988, modified)

The data of both formats must be converted in the molar values to allow comparison with the literature or published results (Buchner & Kiefhaber 2005). Considering the Lambert-Beer law (3.3) the difference of absorption between left A_L and right-handed A_R circularly light can be defined according to equation (3.5):

$$\Delta A = A_L - A_R = \varepsilon_L \cdot c \cdot d - \varepsilon_R \cdot c \cdot d \quad (3.5)$$

For the same sample the concentration c in mol l^{-1} and the path length d in cm for both absorptions are identical. The conclusion is that the difference in absorption correlates directly to the difference of their circular dichroic extinction coefficient ε_L and ε_R . The result is the differential molar circular dichroic extinction coefficient $\Delta\varepsilon$ in $\text{l mol}^{-1} \text{ cm}^{-1}$ which is shown in equation (3.6).

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \quad (3.6)$$

If the CD spectrum is recorded in ellipticity against the wavelength, the results have to be converted by equation (3.7):

$$[\theta] = \frac{\theta \cdot MRW}{d \cdot \beta} \quad (3.7)$$

where θ is the measured ellipticity in millidegrees, MRW is the mean residue weight in g mol^{-1} , d is the path length of the cuvette in mm, β the protein concentration in mg ml^{-1} and $[\theta]$ is the mean residue ellipticity in $\text{deg cm}^2 \text{ dmol}^{-1}$. The mean residue weight is the molecular weight divided by the number of peptide bonds (= number of amino acids minus one, except the protein is acetylated). The interrelation between the two formats is given in equation (3.8) (Greenfield 2006):

$$\Delta\varepsilon = [\theta] / 3298 \quad (3.8)$$

As a result of this work, only spectra in the far UV were recorded (180 – 260 nm). In this region (Figure 3.14) the most significant and consistent signals are from α - helices which show minima near 222 and 208 nm and a maximum near 195 nm. The signals of β - sheets are weaker and more dissimilar because factors, like the length and orientation as well as the twist of the sheets, have an influence on the shape of the CD spectrum. Nevertheless, there is a trend that the signals are positive between 190 to 200 nm and negative from 212 nm to 225 nm. In proteins with the same percentage of α - helices and β – sheets the signals of the helices dominate in the spectra. Loops, random coil and unfolded proteins show a weak signal at 210 nm and a huge minimum between 195 and

200 nm. The conclusion of these properties is that the signal near 195 nm indicates if the protein is in a folded or an unfolded state. Theoretically, this region could be used to monitor the transition during the protein folding process. In practice, the problem is that most buffers have a strong absorption below 200 nm and the results would not be reliable (Buchner & Kiefhaber 2005). The results of the CD measurements are discussed in chapter 4.1.5 on page 75.

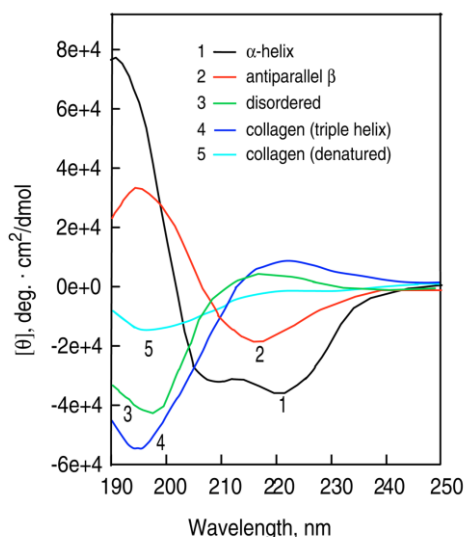


Figure 3.14: CD – spectra of typical secondary structures in the far UV; 1) poly-L-lysine at pH 11,1 α -helical and; 2) in the antiparallel β sheet conformation; 3) poly-L-lysine at pH 5,7 unfolded conformation; 4) collagen in its native and 5) in its denatured form (Greenfield 2006)

3.1.3.3 Fluorescence spectroscopy

Fluorescence spectroscopy has turned out to be a very useful tool in several scientific disciplines. This method is highly sensitive and easier to handle than radioactive tracers. Additionally, it is also possible to reveal information on the single-molecule level (Lakowicz 2006).

The principle of this technique is based on the emission of light. Generally, this phenomenon is called luminescence which can be divided into fluorescence and phosphorescence. Substances which emit light by fluorescence are called fluorophores. The process of fluorescence requires the excitation of electrons from a π -electron system. Consequently, most fluorophores are aromatic or heterocyclic. Initially, one of

two antiparallel orientated electrons of the binding π -orbital gets excited to the anti-binding π^* -orbital by absorbing a photon (Figure 3.9) (Galla 1988). The detailed mechanisms of the absorption and the following transitions can be described by a so called Jablonsky diagram (Figure 3.15).

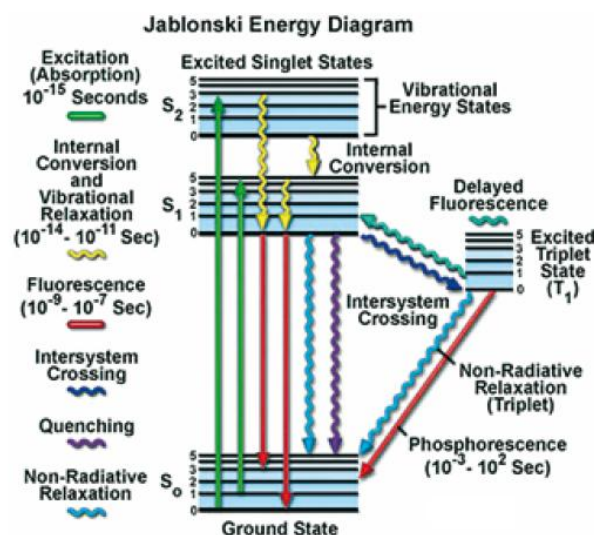


Figure 3.15: Jablonski diagram of different excited states and possible transitions (Wegerhoff et al. 2007)

According to the diagram the electrons are able to occupy different energetic states (S_0 , S_1 , S_2 , ...). Furthermore, at each state the fluorophore can exist in different vibrational sub-states (0, 1, 2, ...). Usually, at room temperature the electrons are in thermal equilibration and occupy the lowest vibrational ground state. Transitions are illustrated vertically in accordance to the Franck-Condon principle. It claims that nuclei displacement is not possible during excitation and emission. The energetic states S_0 and S_1 are too far apart to induce fluorescence by heat. Therefore, an electron gets excited at the excitation wavelength by a photon within 10^{-15} s to a higher singlet state (S_1 or S_2) as well as to a higher vibrational state. After the absorption the excited electron drops down to the lowest vibrational state 0 of the excited singlet state S_1 . This phenomenon is defined as internal conversion. It happens within 10^{-12} s due to thermal equilibration and before any emissions. At this point, the excited electron can return back along several different ways. If the antiparallel orientated electron emits a photon within 10^{-8} s and gets back to the ground state, the transition is called fluorescence. Besides, the

fluorescence intensity can be decreased by several alternative transitions like non-radiative decay, quenching, solvent relaxation or resonance energy transfer (p. 45) (Lakowicz 2006). For this work the emission spectra of the fluorophores Alexa488, Alexa647 and the FRET pair were recorded by the fluorescence spectrometer Perkin-Elmer LS50B whose setup is illustrated in Figure 3.16. Therefore, the sample concentration was chosen with an optical density below 0.1. Generally, light from the light source encounters the excitation monochromator to generate the desired wavelengths. Then, the light beam gets split up to measure one part as reference and the other part enters the sample solution. The emitted light is detected in a 90° angle. Before the signal gets amplified by the sample photomultiplier the light passes the emission monochromator to isolate the fluorescence wavelength. Additionally, slits and filters can be used to adjust the wavelength interval of the excitation and emission light. With this configuration of the spectrometer an emission spectrum can be recorded by measuring the wavelength distribution of the emission at a single excitation wavelength. The output of the uncorrected spectrum gets plotted as fluorescence intensity against wavelength in nm. The emission spectrum of Alexa488 which is displayed in Figure 3.17 represents a typical result of a fluorescence spectrometer (Lakowicz 2006).

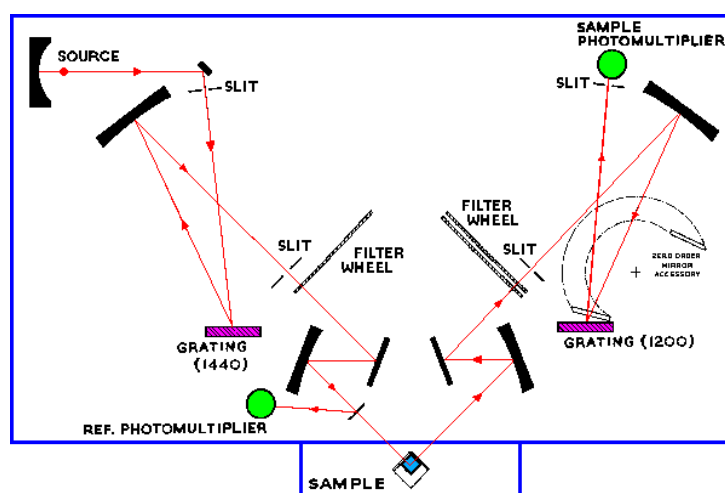


Figure 3.16: Schematic depiction of the Perkin-Elmer LS50B fluorescence spectrometer (Hornak 1999)

Such a fluorescence emission spectrum is usually characterised by three phenomena of fluorescence. The first one is the Stokes shift which was discovered by Sir George Gabriel Stokes. He discovered that a fluorophore emits light at higher wavelengths in

comparison to its excitation wavelength. The reason for higher wavelengths is because of energy losses due to internal conversion after the absorption. Furthermore, energy is lost by thermal effects because the electron reaches a higher vibrational ground state after the transition. These events are also displayed in the Jablonski diagram (Figure 3.15). Besides, there are other effects which increase the Stokes shift like solvent interactions or energy transfer. Consequently, less energy is directly proportional to a higher wavelength according to equation (3.2). The second principle is called Kasha's rule. It states that the shape of the emission spectrum does not depend on a single excitation wavelength. The rule is based on the fact that all fluorescence emissions originate from the lowest vibrational state of S_1 . In addition, this state is reached at different excitation wavelengths within 10^{-12} s by internal conversion. Consequently, this time frame is too short to notice any differences in the emission spectrum. The third principle is called the mirror image rule. The conclusion of this rule is that the emission spectrum represents the mirror image of the $S_0 \rightarrow S_1$ absorptions of the fluorophore (Lakowicz 2006). This image (Figure 3.17) appears because the absorption spectrum is characterised by the vibrational levels of the excited states and the emission spectrum by the vibrational levels of the ground state. If both vibrational levels are similar the image of both spectra seems to have a symmetric axis (Galla 1988). Exceptions of this rule include the pH dependency of the fluorophore with different protonated states or the formation of a charge-transfer complex (Lakowicz 2006).

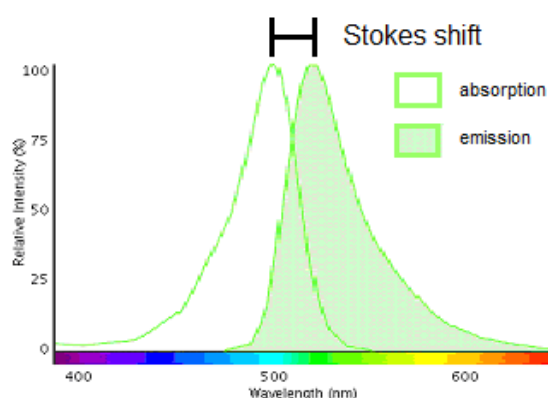


Figure 3.17: Absorption and emission spectrum of Alexa 488 (Life Technology Corporation 2013a, modified)

Practically, the fluorophores can be characterised by their fluorescence lifetime and quantum yield. The fluorescence lifetime is defined as the average time the fluorophore exists in the excited state before it returns to the ground state. Hence, the quantum yield is defined as the ratio of the number of emitted photons to the number of absorbed photons. Both reference marks depend on the emission rate of the fluorophore Γ and the sum of all non-radiative decays k_{nr} . Thus, the fluorescence lifetime τ in ns can be calculated according to equation (3.9) and the quantum yield Q in accordance to equation (3.10) (Lakowicz 2006).

$$\tau = \frac{1}{\Gamma + k_{nr}} \quad (3.9)$$

$$Q = \frac{\text{emitted photons}}{\text{absorbed photons}} = \frac{\Gamma}{\Gamma + k_{nr}} \leq 1 \quad (3.10)$$

Furthermore, fluorophores can be categorized into intrinsic and extrinsic fluorophores. Intrinsic fluorophores are part of the sample like aromatic amino acids in proteins and extrinsic ones are attached to non-fluorescent samples. Tryptophan is the most important intrinsic fluorophore for protein unfolding transitions. The indole group of this amino acid absorb in the near UV at 280nm and usually emit light at 340nm. The emission depends on the chemical environment surrounding tryptophan. A hypsochromic shift occurs, if the amino acid is buried in the protein. But if the indole group is exposed to the solvent, the emission spectrum results with a bathochromic shift (Lakowicz 2006). Therefore, the emission spectrum of tryptophan elucidates whether or not the protein is in a folded or unfolded state. This method was utilized by Osváth et al. (2005) to analyse the folding properties of the N- and C-domain of yeast phosphoglycerate kinase. Single-tryptophan mutants were created for the N- domain, C- domain and the whole protein. The unfolding and refolding transitions were induced by the denaturant guanidine hydrochlorid via stopped-flow techniques and the kinetics were monitored. As a result, the refolding transitions for both domains include an intermediate state and appear to be asymmetric. Finally, this work revealed that domain-domain interactions are the reason why the C- domain folds faster than the N- domain.

3.1.3.4 Fluorescence lifetime measurements

As mentioned before, the fluorescence lifetime τ is the average time of the fluorophore in the excited state. More precisely, the average value of a single exponential decay is the result of 63 % of the molecules with shorter and 37 % with larger decay rates. The fluorescence lifetime can be determined by the time domain method. Therefore, the sample gets excited with a pulse of light. Hence, the pulse width must be significantly shorter than the lifetime of the sample. Additionally, a polarizer is placed in the emission beam at the magic angle ($54,7^\circ$ from the vertical z-axis) to avoid depolarisation effects. One technique of the time domain method, which was utilized for this work, is the so called time correlated single photon counting (TCSPC). TCSPC requires a fluorescence spectrometer with a light source like a picosecond diode laser and a microchannel plate photomultiplier tube (MC-PMT) as high speed detection system. The laser is necessary to create light pulses with small pulse widths. The principle of this technique is based on this special equipment which allows the detection of only one photon per 100 laser pulses. Consequently, it is possible to measure the time between the excitation and detection of one photon with the adequate electronics. The results are summarized in a histogram where photon counts of certain time differences are plotted against the time difference. Multiple counts give a histogram with a form illustrated in Figure 3.18 (Lakowicz 2006).

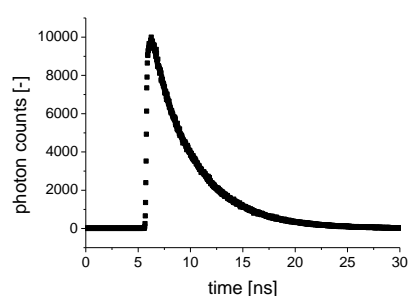


Figure 3.18: Fluorescence lifetime measurement of single labelled myPGK with Alexa488 at 25,5 °C

If the recorded decay is single exponential, it behaves according to equation (3.11) (Lakowicz 2006).

$$I(t) = I_0 \cdot e^{-\frac{t}{\tau}} \quad (3.11)$$

This formula expresses the intensity decay of the starting intensity I_0 ($t = 0$) which drops exponentially with time t . The lifetime τ represents the slope of the $\log(I_0)$ function against time. More complex decays can be calculated by taking the fraction α (<1) of every partial lifetime into account according to equation (3.12) (Lakowicz 2006).

$$I(t) = I_0 \cdot \sum_i \alpha_i \cdot e^{-\frac{t}{\tau_i}} \quad (3.12)$$

3.1.3.5 Förster resonance energy transfer - FRET

FRET, also known as Fluorescence resonance energy transfer, is a nonradiative process which reduces the fluorescence intensity of an excited fluorophore (donor) by transferring the excitation energy to another molecule (acceptor). In contrast to quenching processes the resonance energy transfer is defined as long range dipole-dipole interaction. If the acceptor is also a fluorophore, the interaction results in a so called sensitized fluorescence. The energy transfer includes the following steps: donor excitation, dipole-dipole interaction between the donor and the acceptor and emission from the acceptor (Figure 3.19) (Galla 1988).

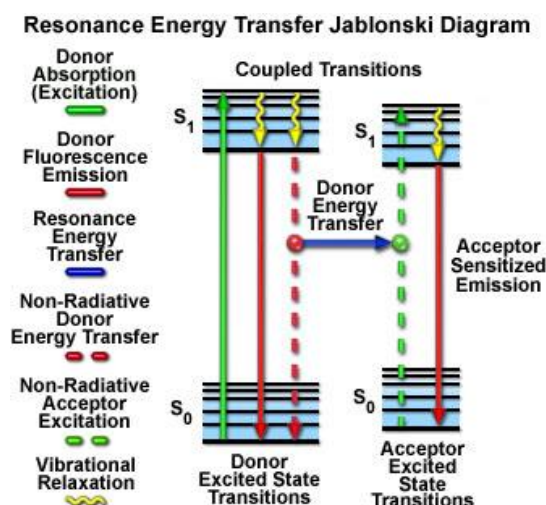


Figure 3.19: Jablonski diagram of the Förster resonance energy transfer (Herman et al. 2012, modified)

The mechanism of the energy transfer happens without the appearance of a photon because it involves a resonance process according to Förster. The electronic system of the donor has the same properties as a mechanic oscillator. Its excitation energy (donor) gets transferred to a second oscillator (acceptor) in the same manner as with coupled pendulums. The transfer rate of the energy k_T in ns^{-1} from the donor to the acceptor correlates with their distance r in Å according to equation (3.13):

$$k_T(r) = \frac{1}{\tau_D} \left(\frac{R_0}{r} \right)^6 \quad (3.13)$$

where τ_D is the lifetime of the donor in ns in absence of the acceptor and R_0 is the Förster radius in Å. The Förster radius is a specific distance for each donor-acceptor pair. It is defined as distance at which the efficiency of energy transfer EET is 50 %. The distance dependence of this method is ideal to acquire structural information on a molecular level (20 - 90 Å). Every change in distance results in a change of the efficiency of energy transfer which can be quantified (Lakowicz 2006). Thus, this technique was utilized for this work to measure the temperature dependent distance between the two domains of modified yeast phosphoglycerate kinase (Figure 3.20). Hence, the protein was labelled with the extrinsic donor Alexa488 and the extrinsic acceptor Alexa647 at the position 155 and 310 (chapter 4.1.4). The steady state data of the protein assemble was obtained for every temperature point.

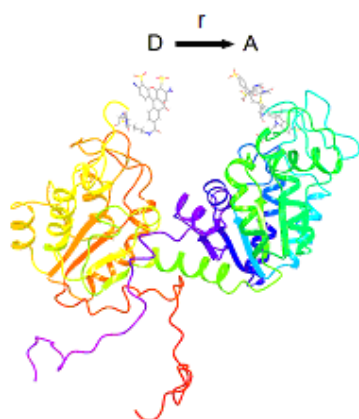


Figure 3.20: Energy transfer between the donor D (Alexa 488) and the acceptor A (Alexa 647) to determine the temperature dependent distance r of modified yeast phosphoglycerate kinase (Desmond 2012)

Practically, the transfer rate $k_T(r)$ results from the spectral properties of the donor and acceptor and can be calculated by equation (3.14).

$$k_T(r) = \frac{Q_D \kappa^2}{\tau_D r^6} \left(\frac{9000(\ln 10)}{128 \pi^5 N_A n^4} \right) J(\lambda) \quad (3.14)$$

Therefore, the transfer rate also depends on the quantum yield Q_D of the donor without the presence of the acceptor, the orientation factor κ^2 of the fluorophores transition dipoles, the refractive index of the surrounding solvent n and the spectral overlap $J(\lambda)$ of the donor emission and the acceptor absorption (Figure 3.21). Furthermore, N_A is the Avogadro's constant or Loschmidt constant which is defined as $6,022141 \cdot 10^{23}$ particles mol^{-1} (Lakowicz 2006).

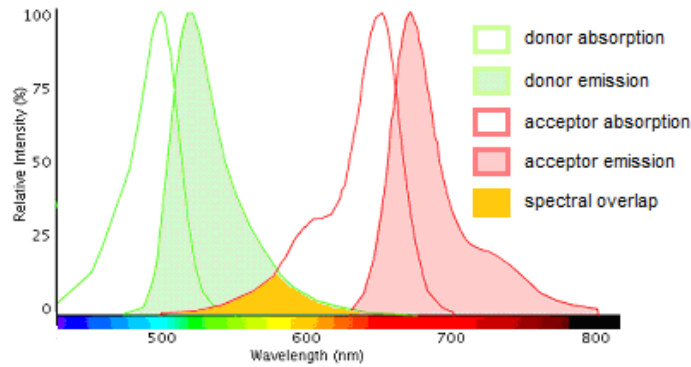


Figure 3.21: Spectral overlap $J(\lambda)$ of the emission spectrum of the donor Alexa488 and the absorption spectrum of Alexa647 (Life Technology Corporation 2013a, modified)

The area of the spectral overlap $J(\lambda)$ depicted in Figure 3.21 is determined by integration according to equation (3.15).

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda}{\int_0^\infty F_D(\lambda) d\lambda} \quad (3.15)$$

The integral consists of the corrected donor fluorescence intensity F_D from λ to $\lambda + \Delta\lambda$ whose total intensity is normalised to unity and the extinction coefficient of the acceptor ε_A of each wavelength (λ). If ε_A is expressed in $\text{M}^{-1} \text{cm}^{-1}$ and the wavelength in nm, the spectral overlap results in the unit $\text{M}^{-1} \text{cm}^{-1} \text{nm}^4$ (Lakowicz 2006).

Unfortunately, the transfer rate k_T is a function of distance. Therefore, it is not suitable to calculate the distance of a FRET pair. In order to quantify r the Förster radius R_0 is determined at first. At the Förster distance the transfer rate is equal to the lifetime of the donor ($k_T = \tau_D$). Consequently, the equations (3.13) and (3.14) can be combined to calculate the Förster distance in Å according to equation (3.16).

$$R_0 = 0,211(\kappa^2 n^{-4} Q_D J(\lambda))^{1/6} \quad (3.16)$$

Usually, the orientation factor κ^2 for an ensemble measurement is assumed to be 2/3 because of isotropic distribution of the transition dipoles. Hence, the extrinsic fluorophores are covalently labelled to the protein. Under this circumstance it was assumed that κ^2 deviates from 2/3. Unfortunately, it is not possible to measure experimentally the orientation factor (Lakowicz 2006). Consequently, simulations were used to determine the temperature dependent values of κ^2 . Further details about the methodology and the results are discussed in the simulation chapter on page 54 and 88.

In the next step, the total fluorescence intensities are measured of the donor F_{DA} with and without the presence of the acceptor F_D . The ratio of these intensities is used in equation (3.17) to calculate the efficiency of the energy transfer EET .

$$EET = 1 - \frac{F_{DA}}{F_D} \quad (3.17)$$

Besides, it is also possible to determine the efficiency by measuring the lifetimes of the donor τ_{DA} with and without the presence of the acceptor τ_D according to equation (3.18). If the measured lifetimes follow a multi-exponential decay, the averaged values are used by summarizing the $\alpha_i \tau_i$ products. In this case, α_i is the relative percentage of the lifetime τ_i .

$$EET = 1 - \frac{\tau_{DA}}{\tau_D} \quad (3.18)$$

The efficiency EET correlates with the fraction of excitation energy of the donor that gets transferred to the acceptor. This value also strongly depends on the distance according to equation (3.19).

$$EET = \frac{R_0^6}{R_0^6 + r^6} \quad (3.19)$$

Finally, the distance between the FRET pair can be calculated after the determination of the Förster radius and the efficiency of energy transfer. Theodor Förster was the first who realized that the energy transfer depends on $1/r^6$ (Förster 1948). This dependence is illustrated in Figure 3.22.

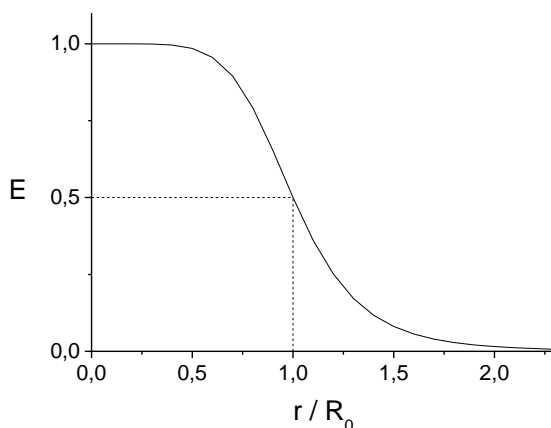


Figure 3.22: Distance dependency of the energy transfer efficiency E , R_0 : Förster distance (Lakowicz 2006)

The consequence of this strong correlation is that relatively small deviations from the Förster distance result in huge efficiency changes. On one hand, if r is equal to $0,5 R_0$, the resulting efficiency is at 0,985. On the other hand, if r is equal to $2 R_0$, the efficiency is 0,015. Consequently, at smaller and larger distances it is not reliable to detect any significant changes in efficiency. Therefore, the realistic distance measurement range for FRET with a Förster distance about 50 Å is between 20 to 90 Å (Lakowicz 2006).

3.1.4 Calorimetric method

3.1.4.1 Differential Scanning Calorimetry (DSC)

DSC is the method of choice to validate the structural stability of a protein by thermal denaturation. Besides, it provides thermodynamic parameters of the unfolding process. The principle is to raise the temperature of a sample and reference equally at constant pressure. During the unfolding process the heat flux for the sample cell must be increased to avoid a temperature difference. This can be achieved by a DSC and a typical assembly is shown in Figure 3.23. The technical requirements for this instrument are extreme sensitivity, reproducibility and signal stability. These parameters must be fulfilled because the protein concentration should be below 1 mg ml^{-1} to reduce intermolecular interactions and usually only small sample volumes are available (Buchner & Kiefhaber 2005).

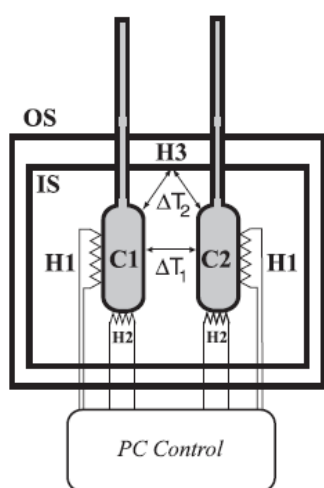


Figure 3.23: Depiction of a DSC instrument

The sample cell (C1) and the reference cell (C2) have identical volumes and are surrounded by the inner (IS) and the outer shield (OS). They are also attached to inlet capillary tubes for filling and cleaning. The main heating elements (H1) supply the cells with constant energy for the temperature upscan. The auxiliary heat elements (H2) are activated to keep the cells at the same temperature ($\Delta T_1 = 0 \text{ } ^\circ\text{C}$). Additional heaters (H3) ensure that the temperature between the inner shield and the cells are the same ($\Delta T_2 = 0 \text{ } ^\circ\text{C}$). All heat elements are regulated via PC (Buchner & Kiefhaber 2005).

The DSC measures directly the heat capacity function $C_p(T)$ in $\text{cal mol}^{-1} \text{ } ^\circ\text{C}^{-1}$ of the protein melting at constant pressure. The theoretical result of such a measurement is shown in Figure 3.24. The actual recorded partial heat capacity $C_{p,pr}(T)$ in Figure 3.24-A can be subdivided into the progress heat capacity $C_p^{prg}(T)$ and the excess heat capacity $\langle C_p(T) \rangle^{exc}$ according to equation (3.20).

$$C_{p,pr}(T) = C_p^{prg}(T) + \langle C_p(T) \rangle^{exc} \quad (3.20)$$

$C_p^{prg}(T)$ is also called the chemical baseline and represents the intrinsic heat capacity of the protein. $\langle C_p(T) \rangle^{exc}$ is calculated by subtracting the chemical baseline from the partial heat capacity. It correlates with the heat absorption during protein unfolding as illustrated in Figure 3.24-B.

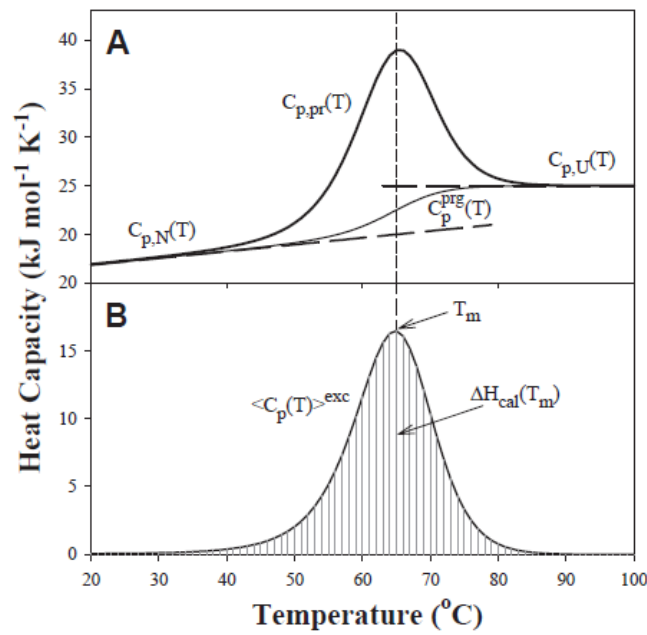


Figure 3.24: Theoretical DSC result of a two state unfolding process

A) DSC upscan; $C_{p,N}(T)$: heat capacity of a protein in the native state, $C_{p,U}(T)$: heat capacity of a protein in the unfolded state, $C_{p,pr}(T)$: partial heat capacity, $C_p^{prg}(T)$: progress heat capacity (= chemical baseline),
 B) evaluated DSC results; $\langle C_p(T) \rangle^{exc}$: excess heat capacity, T_m : melting temperature, $\Delta H_{cal}(T_m)$: calorimetric change of enthalpy (Buchner & Kiefhaber 2005)

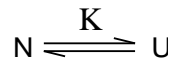
This transition is characterized by the melting temperature T_m and the calorimetric enthalpy $\Delta H_{cal}(T_m)$. $\langle C_p(T) \rangle^{exc}$ reaches its maximum at the melting temperature which correlates with equal protein concentrations in the native and in the unfolded state. The

calorimetric enthalpy results from the area under $\langle C_p(T) \rangle^{exc}$ and is calculated according to equation (3.21).

$$\Delta H_{cal} = \int_0^{\infty} \langle C_p(T) \rangle^{exc} dT \quad (3.21)$$

It represents the model independent enthalpy of the whole unfolding transition. $\Delta H_{cal}(T_m)$ also contains all thermodynamic information of the macroscopic system which can be extracted by the formalism of equilibrium thermodynamics (Buchner & Kiefhaber 2005).

The detailed thermodynamic description of the protein's unfolding transition depends on the unfolding pathway. If the pathway is unknown, it is possible to test whether or not the transition is in accordance with the two-state unfolding model.



The first step to confirm the validity of the possible two-state process requires the calculation of the equilibrium constant $K(T)$ according to equation (3.22). This constant expresses the protein concentration ratio of the unfolded state U to the native state N.

$$K(T) = \frac{[U]}{[N]} = \frac{\int_T^{\infty} \langle C_p(T) \rangle^{exc} dT}{\int_0^T \langle C_p(T) \rangle^{exc} dT} \quad (3.22)$$

The second step includes the calculation of the temperature dependence of the equilibrium constant according to equation (3.23). This dependency is defined by the van't Hoff enthalpy ΔH_{vH} of the two-state model where R is the universal gas constant ($= 8,314 \text{ J mol}^{-1} \text{ K}^{-1}$) and T is the temperature in K.

$$\Delta H_{vH} = -R \cdot \frac{d \ln K(T)}{d(1/T)} \quad (3.23)$$

Finally, if the van't Hoff enthalpy is equal with the calorimetric enthalpy, $\Delta H_{cal}(T_m) = \Delta H_{vH}(T_m)$, the protein unfolding follows a two-state process. However, if the calorimetric enthalpy is larger, $\Delta H_{cal}(T_m) > \Delta H_{vH}(T_m)$, the unfolding process is more complex and includes at least one intermediate state (Buchner & Kiefhaber 2005).

As a result, the temperature dependent description of the simple two-state model requires only three parameters, T_m , $\Delta H(T_m)$ and ΔC_p . These values allow the calculation of all relevant thermodynamic variables. This includes the change of the Gibbs free energy $\Delta G(T)$ in J mol⁻¹, the change of enthalpy $\Delta H(T)$ in J mol⁻¹ and the change of entropy $\Delta S(T)$ in J mol⁻¹ K⁻¹, according to equation (3.24) and the Gibbs- Helmholtz equation (3.25).

$$\Delta G(T) = \frac{T_m - T}{T_m} \cdot \Delta H(T_m) + \Delta C_p \cdot (T - T_m) + T \cdot \Delta C_p \cdot \ln\left(\frac{T_m}{T}\right) \quad (3.24)$$

$$\Delta G(T) = \Delta H(T) - T \cdot \Delta S(T) \quad (3.25)$$

If $\Delta G(T) > 0$, the process is endergonic and thermodynamically not favourable. But if $\Delta G(T) < 0$, the process like protein folding is exergonic and changes until it reaches its minimum. Whether or not a reaction takes place is often temperature dependent because the difference of enthalpy and entropy is often equalising themselves. This phenomenon is called the entropy – enthalpy compensation (Buchner & Kiefhaber 2005). Furthermore, the temperature dependency of both values is defined by the change of heat capacity summarized in equation (3.26).

$$\Delta C_p = \frac{d\Delta H(T)}{dT} = T \cdot \frac{d\Delta S(T)}{dT} \quad (3.26)$$

The protein stability is characterized by ΔC_p , ΔH and ΔS . In theory, the heat capacity change of unfolding ΔC_p results from the exposure of buried groups in the native state to the surrounding solvent. Therefore, large positive contributions arise from the exposure of nonpolar groups and smaller negative ones from polar groups. Another

hypothesis suggests that the disruption of internal interactions has an influence on the heat capacity change. Unfortunately, this could not be validated. Currently, there is not a plausible theory that could entirely predict ΔC_p of an unfolding process. Furthermore, the transition is characterized by the change in enthalpy ΔH . This value is positive because the native state is energetic more favourable than the unfolded one. This effect concludes from unfavourable new interactions with the solvent and the disruption of stabilizing internal interactions. The change in entropy increases during the unfolding process. It is the result of the exposure of buried groups to the solvent and the relative increase of accessible conformations. Generally, the exposed groups entropically favour the native state. This effect gets overcompensated by the increase in conformational freedom and favours the unfolded state. Consequently, entropic effects destabilize the native structure of a protein while larger enthalpic ones preserve it. This contradiction is the reason why proteins are slightly stable. In theory weak protein stability could be evolutionary preferred to guarantee rapid regulation on the protein level. It is also an advantage for the protein degradation by proteases, protein transport including folding-unfolding transitions and to prevent tendencies of aggregation (Buchner & Kiefhaber 2005). Details of the measurements are described in the results chapter (p. 78).

3.2 Simulation methods

3.2.5 Homology Modelling

Homology modelling is a knowledge-based method to predict the tertiary structure of proteins. This technique utilizes the theory of evolutionary related structures within a protein family because of a common ancestor. These protein members share similar sequences which are assumed to have a common basic fold. Therefore, the experimentally determined structure of a protein via X-ray crystallography or NMR spectroscopy can be used as a template for a protein with an unknown structure. Hence, the accuracy of the model depends on the sequence similarity which should be at least 30%. Otherwise, it could result in significant alignment errors (Setlow 2007).

Generally, the homology modelling procedure includes six steps which are displayed in Figure 3.25. First of all, it starts with the template selection which includes the

identification of similar sequences throughout a database. The second step involves the alignment of the target sequence with the template sequence which has the highest similarity. The third step is about the design of a framework structure for the target protein which contains main chain atoms. In the fourth step side chain atoms and loops get added to optimize the model. Afterwards, the whole model is refined and optimized according to energetic parameters in the fifth step. In the sixth and final step the quality of the model gets evaluated (Xiong 2006).

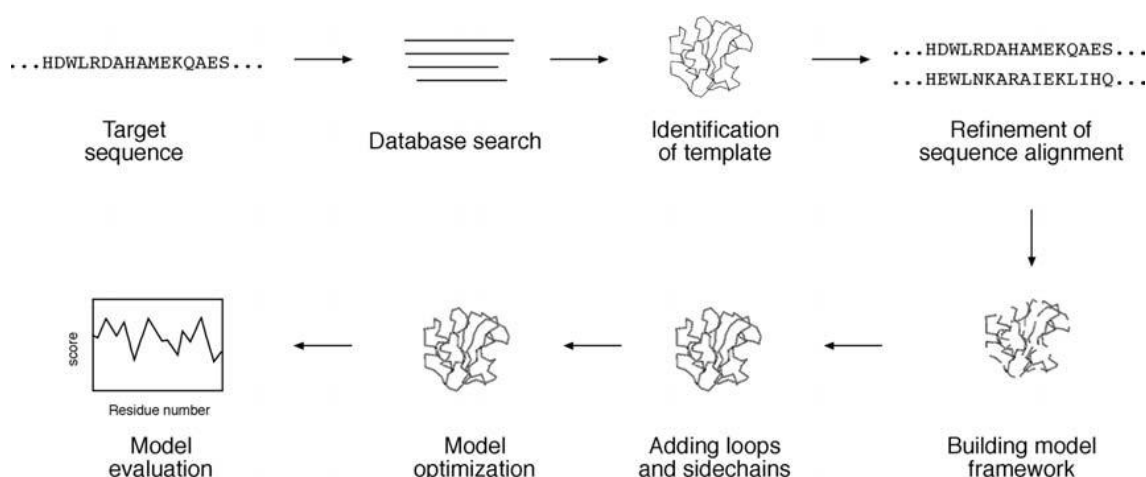


Figure 3.25: Homology modelling procedure (Xiong 2006)

3.2.6 Molecular Dynamics Simulations (MD)

Molecular dynamics is one of many techniques that are part of molecular modelling. The main goal of molecular modelling is a realistic representation of chemical or molecular biological systems. Furthermore, the model is used, especially for new designed and unknown structures, to understand and predict macroscopic properties. The most accurate and realistic prediction of such properties are obtained by the time-dependent Schrödinger equation which is used in quantum mechanics. Unfortunately, this equation can only be applied on relative small molecular clusters to get information at the so called ab-initio level. Therefore, either approximations or empirical parametrizations are used to handle larger and more complex systems in an appropriate time. Generally, molecular dynamics simulations are based on Newton's law of motion according to equation (3.27) (Spoel et al. 2010).

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i, i = 1 \dots N \quad (3.27)$$

This equation describes the force F that acts on a particle with a mass m for each position r and time t . Furthermore, the force correlates with the potential energy function $V(r_1, r_2, \dots, r_N)$ according to equation (3.28) (Spoel et al. 2010).

$$F_i = -\frac{\partial V}{\partial r_i} \quad (3.28)$$

The equations for all particles of a system are solved for each time steps. During the simulation the temperature and pressure is checked to remain at the defined values. The result of these calculations is a trajectory of the system which reaches its equilibrium state after some time. The averaged trajectories are the basis for the extraction of physical properties (Spoel et al. 2010).

The accuracy of the results is limited because of approximations which enables MD simulations of large and complex systems. Deviations from the real values arise because Newton's equations describe the classic mechanics and not the quantum mechanics of atoms. The forces that act on the particles are calculated by force fields which stand apart from the simulation method. They are individually selected and the parameters can be adapted to meet the requirements of the simulation. Hence, the available form of the forces depends on the simulation program (Spoel et al. 2010).

For this work, the MD simulations were performed with the programs GROMACS (GRoningen Machine for Chemical Simulations) (Spoel et al. 2010) and Desmond (2012) of the Schrödinger Maestro software package utilizing the OPLS All-Atom Force Field. The potential energy function V described by this force field is the result of the combined energies from bond stretching, angle bending, non-bond interactions and torsion according to equation (3.29) (Jorgensen et al. 1996).

$$V = E_{bond}(\phi) + E_{angle}(\phi) + E_{n.b.}(\phi) + E_{torsion}(\phi) \quad (3.29)$$

The energy amount for bond stretching and angle bending is acquired by the equations (3.30) and (3.31) which use predefined constants K_θ and K_r (Jorgensen et al. 1996).

$$E_{angle}(\phi) = \sum_{angles} K_{\theta} (\theta - \theta_{eq})^2 \quad (3.30)$$

$$E_{bond}(\phi) = \sum_{bonds} K_r (r - r_{eq})^2 \quad (3.31)$$

The energy calculation for the non-bonded interactions between a particle a and b consists of the Coulomb and Lennard-Jones term described by equation (3.32) (Jorgensen et al. 1996).

$$E_{n.b.}(\phi) = \sum_i^{\text{on a}} \sum_j^{\text{onb}} \left[\frac{q_i q_j e^2}{r_{ij}} + 4\epsilon \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) \right] f_{ij} \quad (3.32)$$

$f_{ij} = 0.5$ for 1,4-interactions; otherwise, $f_{ij} = 1.0$

Usually, the programs use a cut-off radius for those non-bonded interactions. Otherwise, it could happen that the simulations break down because of charged particles in the system. In order to handle such systems, the particle-mesh Ewald (PME) algorithm is used to calculate long-range electrostatics (Spoel et al. 2010).

The last energy term involves the torsional energy E_{torsion} which is defined by equation (3.33). It depends on the dihedral angle ϕ_i , the coefficients for the Fourier series (V_1 , V_2 , V_3) and the phase angles (f_1 , f_2 , f_3) (Jorgensen et al. 1996).

$$E_{torsion}(\phi) = \sum_i \left\{ \frac{V_1^i}{2} [1 + \cos(\phi_i + f_{i1})] + \frac{V_2^i}{2} [1 - \cos(2\phi_i + f_{i2})] + \frac{V_3^i}{2} [1 + \cos(3\phi_i + f_{i3})] \right\} \quad (3.33)$$

In order to perform a simulation the system must be generated and prepared according to a setup described in the results chapter (p. 89). Each setup step was carried out by using different algorithms to meet the required conditions (temperature, pressure, time, ion concentration). The general function of diverse algorithms are discussed below which were used mainly with GROMACS (p. 114) but as well as with Desmond.

3.2.6.1 Leap frog integrator

MD simulations are performed under predefined time steps Δt of a few picoseconds. GROMACS utilizes the leap frog algorithm which calculates the positions and velocities with a time difference of $\frac{1}{2} \Delta t$. Consequently, the calculations of further positions and velocities depend on the results of previous time steps. Therefore, this algorithm is time reversible (Spoel et al. 2010).

3.2.6.2 Periodic boundary condition

A common parameter for a MD simulation is the periodic boundary condition (PBC). PBC allows a particle the re-entrance on the opposite site of the simulation box. Thus, it is possible to simulate the conditions of a bulk system with small errors for large systems (Spoel et al. 2010).

3.2.6.3 Steep descent minimization

Before any simulations gets started the system should be energetic minimized. The principle behind the energy minimization is based on the multidimensional hypersurface of the potential energy function. On this complex theoretical landscape there are many local and one global minimum of the system. After the energy minimization it is usually not possible to distinguish if the system reached a local or the global minimum. With GROMACS the *Steepest decent* method was utilized. It follows the direction of the negative gradient without saving the history of previous steps. This is a fast and simple method which ends in a minimum nearby the starting point on the hypersurface (Spoel et al. 2010).

3.2.6.4 Linear Constraints Solver (LINCS)

GROMACS uses the LINCS algorithm to implement holonomic constraints. The algorithm works with three position updates within every time step. Firstly, the update is performed without the constraints. Secondly, the forces work along the bond. At last, the bond is corrected to its original length after the movement. This method is not iterative which makes it relatively fast and stable. Generally, LINCS proved to be useful to simulate Brownian dynamics (Spoel et al. 2010).

3.2.6.5 Berendsen thermostat

In order to perform a temperature equilibration of a system it is necessary to implement the conditions of a canonical ensemble. This ensemble is defined by a system with constant number of particles, constant volume and constant temperature (NVT). Different algorithms can be used to perform NVT equilibrations but usually the Berendsen thermostat is utilized. The Berendsen algorithm operates like an external heat bath and keeps the predefined temperature of the system T_0 . Deviations are corrected according to a first order kinetic which decays by the time constant τ . The mathematical formalism is described by equation (3.34) (Spoel et al. 2010).

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad (3.34)$$

3.2.6.6 Velocity-rescaling thermostat

The simulations with GROMACS were performed with the velocity-rescaling thermostat. This algorithm is equal to the Berendsen thermostat but with an additional stochastic term. The additional term guarantees that the kinetic energy distribution is correct. Consequently, this algorithm allows the simulation of a reliable canonical ensemble (Spoel et al. 2010).

3.2.6.7 Parrinello-Rahman barostat

In a similar way it is possible to perform a pressure equilibration of a system. Therefore, the condition of an isothermal-isobaric ensemble is generated. This ensemble is defined by a constant number of particles, constant pressure and constant temperature (NPT). For this work the Parrinello-Rahman barostat was utilized. In theory it generates a true NPT ensemble (Spoel et al. 2010).

3.2.7 Determination of the orientation factor - κ^2

The orientation factor κ^2 is a geometric value which influences the transfer rate k_T between the donor and acceptor as described in equation (3.14). It represents the relative orientation of the transition emission dipole of the donor to the transition absorption dipole of the acceptor and is calculated according to equation (3.35) or (3.36).

$$\kappa^2 = (\cos \theta_T - 3 \cos \theta_D \cos \theta_A)^2 \quad (3.35)$$

$$\kappa^2 = (\sin \theta_D \sin \theta_A \cos \phi - 2 \cos \theta_D \cos \theta_A)^2 \quad (3.36)$$

In those equations θ_T represents the angle between both dipoles. θ_D and θ_A are the angles between the dipole of the donor respectively the dipole of the acceptor and the distance vector r . The angle between the planes corresponds with ϕ . The values for this variable can range from 0 to 4. Therefore, if the dipoles are aligned in the same direction as depicted in Figure 3.26, then $\kappa^2 = 4$. Parallel orientations result in $\kappa^2 = 1$ and perpendicular orientated dipoles have not any energy transfer at all (Lakowicz 2006).

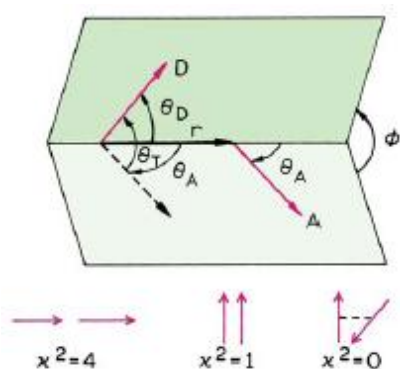


Figure 3.26: The orientation factor κ^2 is the result of the relative direction of the emission dipole of the donor and the absorption dipole of the acceptor (Lakowicz 2006).

Until now, it is hardly possible to determine the orientation factor of a donor-acceptor pair by an experiment. Therefore, κ^2 is assumed to be $2/3$ for isotropic distributed donors and acceptors in a solvent. Furthermore, static donor-acceptor orientations result in a $\kappa^2 = 0,476$. Finally, these assumptions have a maximum error of 35 % because the sixth root of κ^2 is taken to calculate the distance as described by equation (3.16) (Lakowicz 2006).

4 Results

4.1 Experiment

4.1.1 Amplification

The goal of the amplification was to increase the stock of the pET 15b + myPGK vector. Practically this includes four steps: the *transformation*, the *scale up* of the cell preculture, the *purification* with the Miniprep Kit and the *control* by agarose gel electrophoresis.

In the first step, the transformation of E. Coli DH5 α cells with the vector was performed according to the Stanford University (2003) protocol (p. 23). The result was an agar plate with several colonies of transformed cells. Three of these colonies were used in the next step, the scale up. Therefore, a mixture of 10 ml LB medium and 10 μ l of the 0,1 g/ml Carbenicillin solution was inoculated by the selected colonies. The prepared cell suspension was incubated at 37,2 °C and 200 rpm over night. On the next day, the purification was executed with the Miniprep Kit. Consequently, the concentrated cell suspension in the falcon tube was centrifuged for 10 minutes at 4000 rpm and the supernatant was discarded. The pellet was solved in 500 μ l Resuspension™ Solution, mixed with 500 μ l Lysis™ Solution and inverted six times. Then 700 μ l Neutralization™ Solution was added and the tube was inverted again six times. After the falcon tube was spinned for 10 minutes, the supernatant was poured into a 2ml Eppendorf tube and centrifuged for 10 minutes at 16.100 rpm. The supernatant was equally divided into two GeneJet™ spin columns. The spin columns were centrifuged twice for 1 minute after adding 500 μ l Wash™ Solution each time. The solutions were discarded and the empty columns were spinned again for 1 minute. For the elution, the filter elements of the spin columns were transferred to a new tube and filled with 50 μ l Elution™ Buffer. After the incubation time of 2 minutes the columns were spinned for 2 minutes and the filter elements were disposed. The resulting amplified vector solution was stored in two tubes at -20 °C (Fermentas 2006). The concentration of vector

solutions I was 171 ng/ μ l and vector solution II 135 ng/ μ l, determined with Nanodrop and blanked against the Elution™ Buffer.

At last, the successful amplification of the desired nucleotides had to be controlled by agarose gel electrophoresis. Therefore, 50 ml of the 0,7 % agarose gel was made by heating 0,35 g agarose, 1 ml 50x TAE Buffer, 5 μ l SYBR and 49 ml distilled water in the microwave at medium power for 2 minutes. After some shaking, the mixture was heated again for one minute. The hot solution was cooled for a few seconds and poured into the gel form. An eight teeth gel comb was stuck onto the top and the gel stood one hour for polymerisation. The next step was the sample preparation. For that reason the purified BSA (100x) was diluted 1:10 (10x) and the samples were mixed according to Table 4.1.

Table 4.1: DNA sample preparation for agarose gel electrophoresis

pos.	Name	NEBuffer 4	BSA 10x	vector	dH ₂ O	BamH I	Nde I	BPB 6x
		[μ l]						
1	Marker							4
2	undigested vector I	2	2	2	14			4
3	undigested vector II	2	2	2	14			4
4	digested vector I	2	2	2	12	1	1	4
5	digested vector II	2	2	2	12	1	1	4
6	dig. vector I (1:10)	2	2	2	12	1	1	4
7	dig. vector II (1:10)	2	2	2	12	1	1	4
8	Marker							4

Generally the dH₂O, BSA (10x) and NEBuffer 4 got vortexed. The NEBuffer 4 was chosen to match the requirements of both restriction enzymes (BamH I, Nde I). Then the vectors and restriction enzymes were added, the mixture vortexed and put on ice. After 15 seconds of centrifugation the samples stood one and a half hours for digestion. Afterward, 4 μ l bromphenol blue (BPB) came to each sample in order to mark the front. The gel was put into the electrophoresis and covered with TAE buffer. The pockets were filled with 10 μ l sample or 1 μ l marker solution according to Table 4.1. At the end, the electrophoresis ran 31 minutes at 84 V, 0,05 A and 4 W.

The finished gel was illuminated with UV and the outcome is illustrated in Figure 4.1.

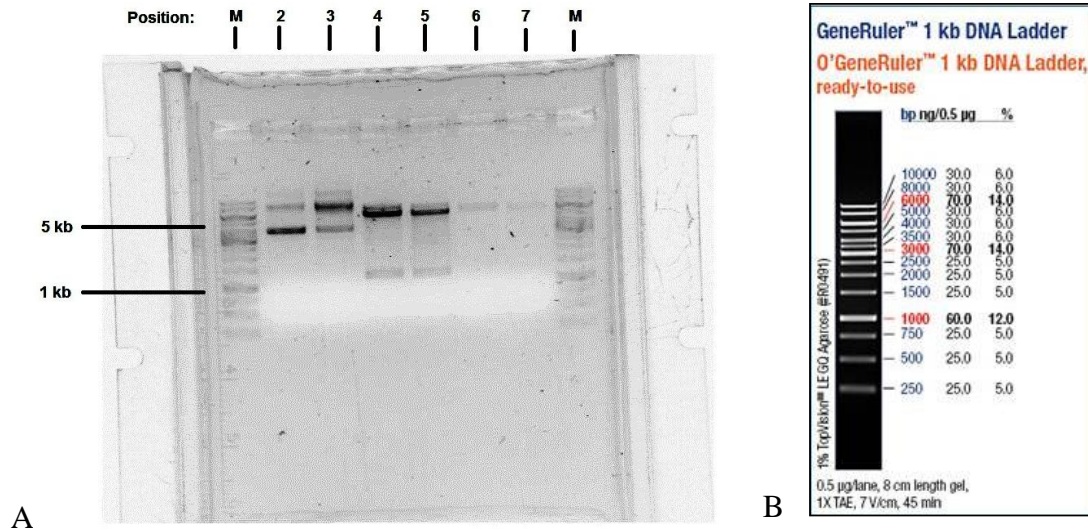


Figure 4.1: Amplification control – A) edited negative picture of the finished DNA electrophoresis, M: marker (GeneRuler™ 1kb DNA Ladder), 2+3: undigested pET15b + myPGK vector, 4+5: digested vector, 6+7: digested vector (1:10); B) marker description (Thermo Scientific 2013a)

The gel shows that the undigested pET15b + myPGK vectors (2; 3) with a size of 6,95 kbp give three bands at 10,30 kbp, 7,60 kbp and 4,13 kbp. This separation occurs because plasmids are able to form different conformations (Figure 4.2). Therefore, different migration velocities are the result of the open, linear and supercoiled state and the determination of the plasmid size would not be reliable. Apart from that, the digested plasmids (4; 5) are fragmented into two parts. The first band appears at 5,92 kbp and correlates with the pET15b vector size of 5,7 kbp. The second one corresponds to the gene of myPGK (1,26 kbp) and stays at 1,20 kbp. These results proof that the amplification of the pET15b + myPGK vector was successful. The amplified vectors were stored at -20 °C after the transformation of the expression cells.

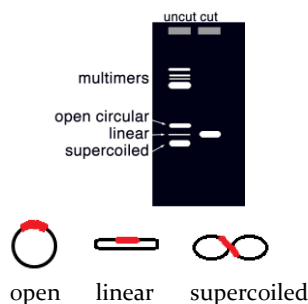


Figure 4.2: Different conformations of plasmids (Rothman 2012, modified)

4.1.2 Expression

After the amplification of the vector, the modified yeast phosphoglycerate kinase was expressed in E. Coli BL21 DE3 cells. This procedure includes the following steps: the *transformation* of the expression cells with the plasmid, the *scale up* of the preculture and the *induction* of the expression with IPTG.

The expression was initiated by transforming the E. Coli BL21 DE3 cells with the pET15b + myPGK vector according to the Stanford University (2003) protocol (p. 23). At the end of the transformation, a selective agar plate was inoculated by the transformed cells and incubated upside down at 37 °C for 23 hours. The rest of the cell culture was stored as glycerol stock for further expression batches at -80 °C. This method is described in the protocol of the Hebrew University of Jerusalem (n.d.) (p. 26). On the next day, three falcon tubes were prepared with 10 ml LB medium and 10 µl of the carbenicillin solution. Two of them were inoculated with one colony forming unit (CFU) of the agar plate and the third one was used as reference. These precultures were incubated at 37 °C and 200 rpm for 18 hours. At the end, the precultures formed successfully milky cell suspensions in comparison to the clear reference. For the scale up, three batches with a selective solution of 500 ml LB medium and 500 µl Carbenicillin solution were mixed with 500 µl of preculture I. According to the work of Rosenkranz (2011a), these cultures were incubated at 37 °C and 150 rpm until the OD (600 nm) values exceeded an absorption of 0,6. This aim was achieved after 6 hours. At this point, the protein expression was initiated with 1 mM IPTG. Finally, the expression was stopped after 3 hours by centrifuging the suspension at 6000 rpm and 4 °C for 30 minutes. The supernatants were discarded and the cell pellets were stored for the protein purification at -20 °C.

4.1.3 Purification

The protein (myPGK) has been extracted from the expression cells after they were disrupted. In addition, the protein was purified by ammonium sulphate precipitation (p. 27) and the IMAC technique (p. 28) according to the work of Dr. Rosenkranz (2011a). The purification steps are summarised in Figure 4.3.

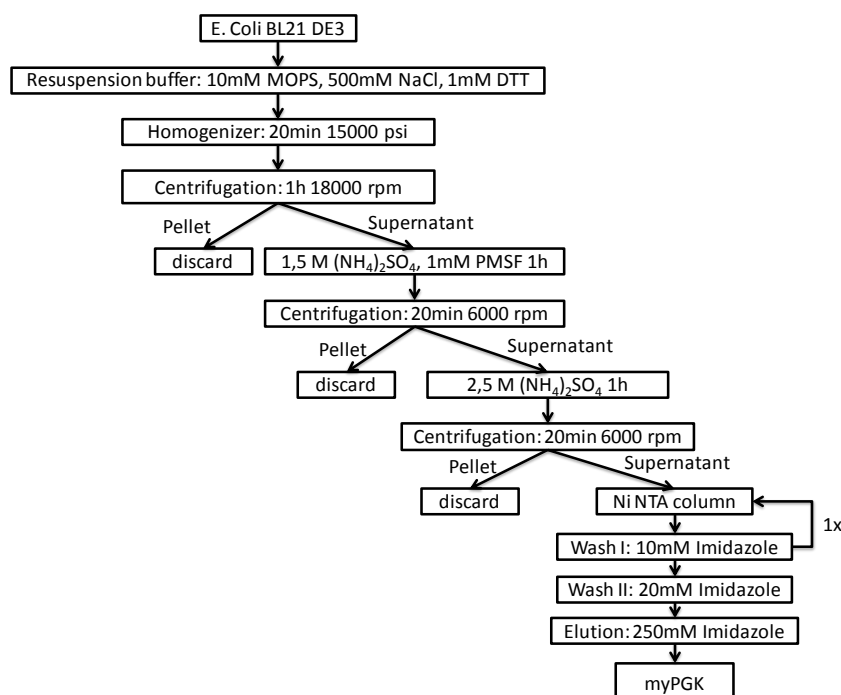


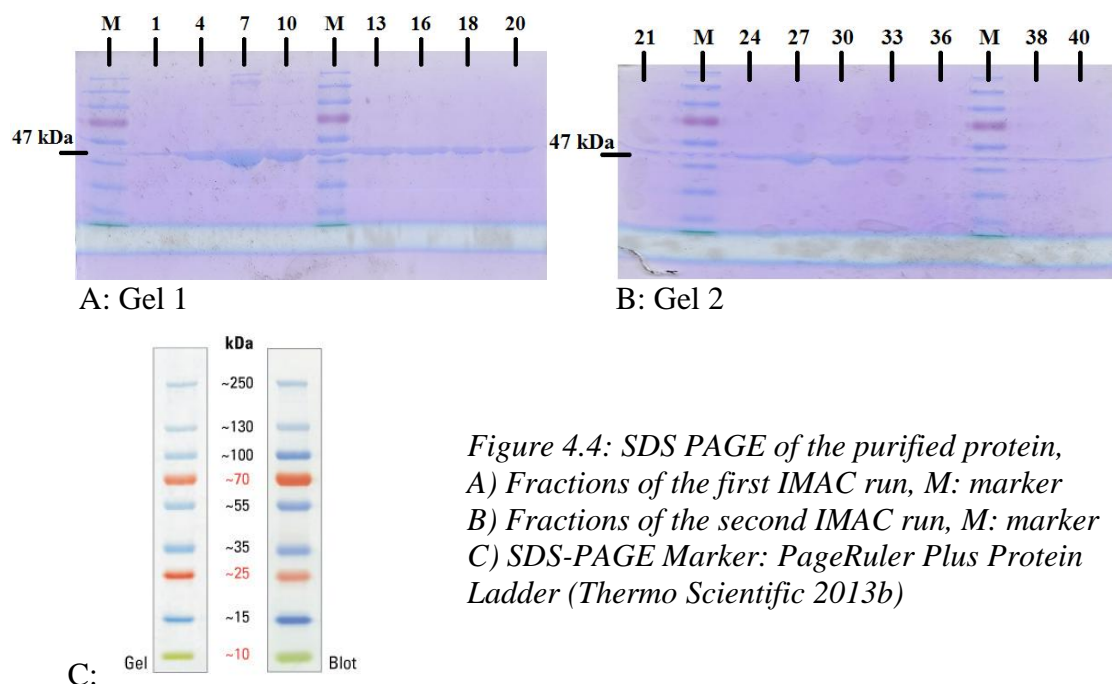
Figure 4.3: Purification procedure of the expressed myPGK in *E. Coli BL21 DE3* cells (Rosenkranz 2011a, modified)

The whole purification procedure was performed at 4 °C to avoid the denaturation of the protein. First of all, the *E. Coli BL21 DE3* cell pellet from the expression was solved in 20 ml Resuspension buffer. In the next step, the cells were disrupted by a dynamic high pressure homogenizer. Therefore, the machine was flushed for 10 min with distilled water before the cells were lysed at 15000 psi for 20 min. At the end, the homogenizer was cleaned with 1 l hot water, 1 l soapy water, water to remove the soap and finally stored in 70 % ethanol. The lysed cell suspension was centrifuged at 18000 rpm for one hour. The pellet was discarded and the supernatant used for the first precipitation step. Therefore, a concentration of 1,5 M ammonium sulphate and 1 mM phenylmethyl-sulfonyl fluoride (PMSF) was established. The serine protease inhibitor, PMSF, was used to eliminate the risk of protein inactivation. The solution was stirred for one hour at 4 °C and centrifuged at 6000 g for 20 min. The pellet was discarded and the supernatant used for the second precipitation step. Therefore, the ammonium sulphate concentration was raised to 2,5 M and the mixture was stirred again for one hour at 4 °C. Afterwards, the solution was centrifuged at 6000 rpm for 20 min and the pellet was discarded. The supernatant was shock frozen with liquid nitrogen and stored at -20 °C over night.

On the next day, two SDS gels were prepared for the electrophoretic analysis of the collected fractions at the end of the purification. The preparation started by solving acrylamid, TRIS and SDS in water with ultrasonic for the running gel. Later the radical initiator ammonium persulphate (APS) and TEMED was added. The gel was instantly poored between two glas plates and covered with isopropyl alcohol to achieve a horizontal surface. After 45 minutes of polymerisation the isopropyl alcohol was discarded and the stacking gel was mixed. The acrylamide, TRIS and SDS was solved again in distilled water by ultrasonic and finally APS and TEMED was added. The gel was instantly poored on top of the running gel and sealed carefully with a ten teeth gel comb to avoid air bubbles.

The purification of the supernatant after the second precipitation step was continued with the 5 ml Ni-NTA column. Hence, the column was flushed with five column volumes (CV) Resuspension buffer (25 ml) at a flow rate of 1ml/min. Then the pre-purified protein solution was applied twice. The column was washed with four CVs Wash I (20 ml) and the fractions were collected for a second run. In addition twelve CVs Wash II (60 ml) were applied. At the end, the protein was eluted with six CVs elution buffer (30 ml) and collected as 1,5 ml fractions. In Dr. Rosenkranz (2011a) work it is mentioned that a significant amount of protein is also in the collected Wash I solution. Therefore, the column was equilibrated with five CVs Wash I (25 ml) before the Wash I fraction of the first procedure was applied twice. Afterwards, the column was washed again with twelve CVs Wash II (60 ml). The protein was eluted with six CVs elution buffer and collected as 1,5 ml fractions.

At the end, the eluted fractions of the first and second purification run were analysed by SDS electrophoresis. For this purpose, 10 µl of every third eluted fraction was mixed with 5 µl Sample buffer, heated at 95 °C for 10 minutes and cooled on ice. The electrophoresis with the SDS gels was filled with Lämmli buffer and the gel pockets with 10 µl of the sample solutions or with 4 µl of the marker (Page Ruler, Plus Prestained Protein Ladder). The gels ran at 120 V, 0,03 A and 4 W for one hour and were stained with the Staining solution for 15 minutes. Afterwards, the gel was shaken three times in Destainer solution and once in water for 10 min. The results of the SDS PAGE are shown in Figure 4.4.



*Figure 4.4: SDS PAGE of the purified protein,
A) Fractions of the first IMAC run, M: marker
B) Fractions of the second IMAC run, M: marker
C) SDS-PAGE Marker: PageRuler Plus Protein
Ladder (Thermo Scientific 2013b)*

According to the outcome of the SDS PAGE gels, the fractions 4 to 20 and 24 to 33 were identified to contain the protein and were combined. Afterwards, 10 ml of the protein solution were dialysed twice against 2 litres Folding Studies Buffer (FB) and 30 ml twice against 2 litres Labelling Buffer (LB). At first, the buffers were exchanged for one hour and at the second time the dialysis ran over night.

In the next step, the protein concentrations were determined according to the Warburg Christian method (p. 34) by measuring the absorption at 260 nm and 280 nm. The protein concentration in LB was 0,428 mg/ml (= 8,80 μ M) and the protein concentration in FB was 0,393 mg/ml (= 8,07 μ M). The sample solutions were fractionated in 2 ml aliquots, shock frozen with liquid nitrogen and stored at -20 °C.

The protein in FB was used for the CD (p. 75), DSC (p. 78) and fluorescence measurements (p. 81). Besides, the protein in LB was utilized for the labelling procedure (p. 68) and the enzyme activity assay (p. 75).

4.1.4 Labelling of myPGK

The labelling procedure was performed according to an optimized standard operation procedure (SOP) from Dr. Rosenkranz. The aim was to obtain a double labelled myPGK (donor + acceptor), a donor only myPGK and an acceptor only myPGK for the FRET measurements. However, at first an ion exchange column had to be packed with diethylaminoethyl- (DEAE) sepharose (= 6 % cross-linked agarose) and evaluated for the labelling procedures. This was necessary to establish the same conditions according to the SOP.

4.1.4.1 Packing a DEAE sepharose column

The packing started by stirring and degassing the packing buffer (10 mM MOPS, 1 M NaCl) under vacuum. Then the buffer was sterilised by using a 0,22 µm filter. The column was filled air bubble free with the prepared buffer and the shipping solution of the DEAE sepharose was decanted. In the next step, the material was suspended with 50 ml packing buffer and the slurry was transferred into a graduated cylinder. In addition the supernatant was removed after one and a half hour of sedimentation. The sediment was resuspended with packing buffer up to a volume of 55 ml and stood again for one and a half hour. At the end, the supernatant was discarded, the sediment was resuspended in 40 ml packing buffer and poured down a glass rod into the column. The slurry stood for two days before the air bubble free column was equilibrated with storage buffer (10 mM MOPS, 20 % ethanol) and placed in the cold store at 4 °C.

4.1.4.2 Evaluation of the DEAE sepharose column

After the packing the column was linked to the FPLC system (Figure 3.6) for the evaluation. This included the equilibration of the system with double distilled water, the injection of 0,2 ml 1 % acetone (= 1 % of the column volume) and the measurement of the absorption at 280 nm. The quality of the packing can be determined by calculating reference values like the height equivalent to a theoretical plate *HETP* in cm by equation (4.1) and the asymmetry factor A_s by equation (4.3):

$$HETP = \frac{L}{N_{th}} \quad (4.1)$$

$$N_{th} = 5,54 \cdot \left(\frac{V_R}{w_h} \right)^2 \quad (4.2)$$

$$A_s = b/a \quad (4.3)$$

where L is the bed height in cm, N_{th} is the number of theoretical plates, V_R is the retention volume in ml, w_h is the peak width at half peak height in ml, a is the first half peak width at 10 % of the peak height in ml and b is the second half peak width at 10 % of the peak height (GE Healthcare n.d.b).

These values are specific for every column and should be measured every time under the same conditions to gain comparable results. This is helpful to check if the performance remains the same. The outcome of the evaluated DEAE column is shown in Figure 4.5.

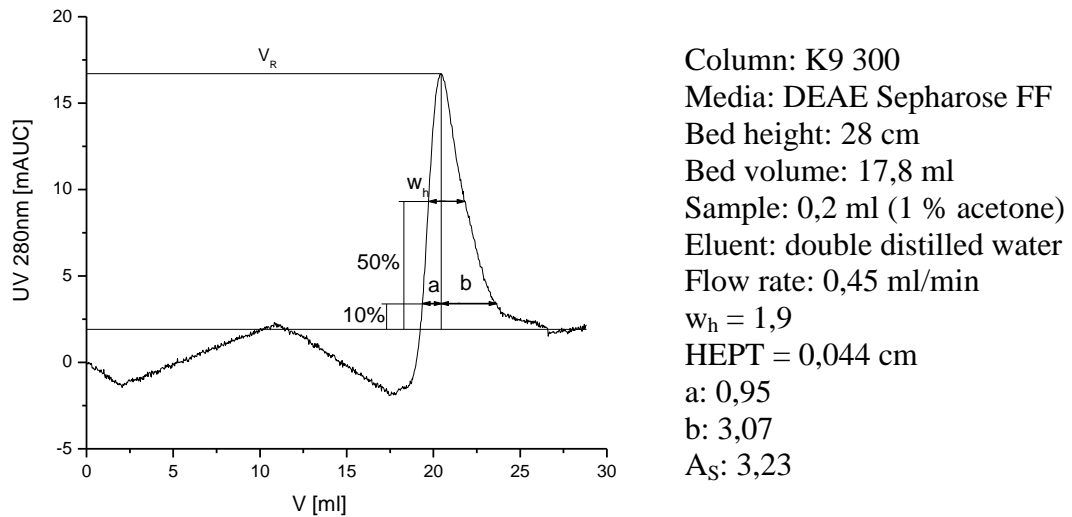


Figure 4.5: Evaluation of the DEAE Sepharose Fast Flow column (0,2ml 1% acetone)

According to the instructions of GE Healthcare (n.d.b) the asymmetric factor should be < 3 . Nevertheless, the performance of the column was sufficient for the labelling (Figure 4.6) procedure although the A_s value was greater than 3 (Figure 4.5).

4.1.4.3 Labelling procedure for the double labelled myPGK

After the evaluation of the column, the purified myPGK was labelled with one acceptor and one donor. This included the following steps: *labelling with the acceptor, purifying the single labelled myPGK, labelling with the donor and the purification of the double labelled protein*. This modus operandi guarantees that only correct double labelled proteins with one donor and one acceptor give a FRET signal. However, it is not possible to distinguish which labelling position (135 or 290) is occupied by the donor and which by the acceptor. Therefore, it is assumed that the two possible double labelled myPGKs (mypgkdl) give equal FRET results.

The labelling procedure was performed within three days. On the first one, 10ml myPGK in LB was centrifuged at 4 °C using a centrifugation concentrator (10.000 MWCO). The concentration increased from 0,393 mg/ml to 2,14 mg/ml (44 µM). In the next step, 30 ml phosphate buffer (100 mM KH₂PO₄, 500 mM NaCl) and 1 ml 1 M NaOH were degassed under vacuum for 30 minutes to remove the interacting oxygen. Additionally, the solutions were bubbled with nitrogen for 15 minutes quite strongly. Then, 3 mg TCEP (reducing agent) were solved in 800 µl phosphate buffer and 200 µl 1 M NaOH which gave a 10 mM solution. Separately, 3 mg glutathione were prepared and stored at 4 °C and one of two PD 10 desalting columns was equilibrated with MOPS buffer (10 mM MOPS) for the next day.

The labelling reaction was initiated by mixing the components of Table 4.2 in an eppi. The head space was blown with nitrogen before sealing and the mixture got vortexed for one minute. The reaction proceeded in the dark at room temperature for 12 hours.

Table 4.2: Composition of the acceptor labelling mixture

component	concentration	note
myPGK	10 µM	
TCEP	200 µM	10 x the number of cysteines, double mutant = 20 x the protein concentration
Alexa647-maleimide	30 µM	3 x the protein concentration
Phosphate buffer		up to 1 ml

On the second day, glutathione was dissolved in 1 ml water (10 mM). Then, the labelling reaction was terminated with 300 µM glutathione (10 times the dye concentration) (Figure 3.8). One of two PD 10 desalting columns was equilibrated with

phosphate buffer before the labelling mixture was applied. Consequently, twenty 0,5 ml fractions were collected and the absorptions at 260 nm, 280 nm and 647 nm were determined. As a result, the fractions with the protein (3-6) were combined and applied to the MOPS buffer equilibrated PD 10 desalting column. After the exchange of the buffer, twenty 0,5 ml fractions were collected and the absorptions were measured again. The fractions 1-5 were combined which contained non-labelled, single labelled and double labelled protein in MOPS buffer.

In the next step, the DEAE Sepharose column was connected to the FLPC system (Figure 3.6) and equilibrated with MOPS buffer at 4 °C. The collected fractions were injected into the system (ÄKTA) via a 5 ml loop. The separation of the different labelled proteins was achieved by a stepwise gradient with the cleaning buffer (10 mM MOPS, 1 M NaCl). Therefore, the system was flushed with two column volumes (40 ml) MOPS buffer to remove the unbound protein from the DEAE column. Then, two CVs (40 ml) with 4 % cleaning buffer were used to elute non-labelled protein. Two CVs (40 ml) of 11,5 % cleaning buffer eluted the single labelled protein. The double labelled protein was removed with two CVs (40 ml) of 18,5 % cleaning buffer. Finally, four CVs (80 ml) of 100 % cleaning buffer were used to clean the column. After the column, several parameters were measured like absorptions at 260 nm, 280 nm and 647 nm, conductivity and pH. At the end, the eluate was fractionated. The results of the separation are illustrated in Figure 4.6.

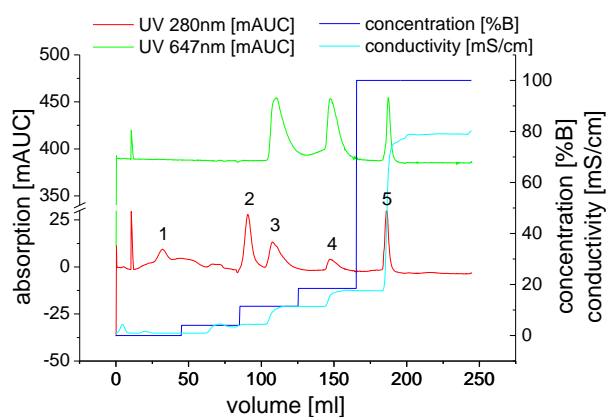


Figure 4.6: Separation of acceptor labelled proteins with DEAE ionexchange FPLC using a step gradient (B: cleaning buffer), 1: unbound protein, 2: non-labelled protein, 3: single labelled protein, 4: double labelled protein, 5: waste

The fractions (101 - 130 ml) of peak 3 (Figure 4.6) with the single labelled protein were combined and concentrated to 2,5 ml using a 10.000 MWCO centrifugation concentrator. Therefore, the first half was spun at 4000 g for 10 minutes. Then, the second half was added and spun again for 10 minutes at 4 °C. The concentrate was applied to the phosphate equilibrated PD-10 desalting column. After the exchange of buffer, twenty 0,5 ml fractions were collected and the absorptions at 260 nm, 280 nm and 647 nm were measured. The fractions 2-7 with the labelled protein were combined and concentrated for 5 minutes at 4000 g with a 10.000 MWCO concentrator to 1 ml. The protein concentration was determined by measuring the absorption of the acceptor at 647 nm. As a result 1,49 μ M of the protein was single labelled with the acceptor which correlates to a yield of 14,9 %.

At this point, the second labelling procedure was started with the donor (Alexa488). Therefore, 30 ml phosphate buffer, 1 ml 1 M NaOH and the single labelled protein solution were degassed in vacuum for 30 minutes. Then, the solutions were bubbled with nitrogen for 15 minutes. After that, a 10 mM TCEP solution was made by solving 3 mg TCEP in 800 μ l phosphate buffer and 200 μ l 1 M NaOH. Furthermore, 3 mg glutathione was prepared and stored at 4 °C for the next day. Finally, the labelling reaction with the donor was initiated by mixing the components according to Table 4.3 in an eppi. The headspace was blown with nitrogen before sealing and the mixture was vortexed for a minute. The reaction was performed at room temperature in the dark for 13 hour.

Table 4.3: Composition of the donor labelling mixture

components	concentration	note
Acceptor labelled myPGK	1,49 μ M	
TCEP	30 μ M	10 x the number of cysteines, double mutant: 20 x the protein concentration
Alexa488 maleimide	15 μ M	10 x the protein concentration
Phosphate buffer		Up to 1 ml

On the third day, the prepared glutathione was solved in 1 ml distilled water (10 mM) and 200 μ M was used to terminate the labelling reaction. The protein solution was applied to the phosphate buffer equilibrated PD-10 desalting column. Twenty 0,5 ml fractions were collected and the absorptions at 260 nm, 280 nm and 647 nm were

determined. The fractions 4-8 with the double labelled protein were combined and applied to the MOPS buffer equilibrated PD-10 desalting column. The fractions 2-7 of the twenty 0,5 ml eluates were combined after measuring the absorptions. At the end, the double labelled protein solution was concentrated with a 10.000 MWCO concentrator to 1 ml. The resulting protein concentration of 0,93 μM was determined by measuring the absorption of the acceptor at 647 nm. The protein solution was mixed with 1 μM PMSF, divided into 100 μl fractions, shock frozen with liquid nitrogen and stored at -20 °C.

The labelling efficiency of the acceptor was 28,5 % and of the donor 22,1 % which were calculated according to equation (4.4) (Thermo Scientific 2011):

$$\text{labelling efficiency} = \frac{\text{moles dye}}{\text{moles protein}} = \frac{A_{\text{dye}} \cdot \epsilon_{\text{protein}}}{\epsilon_{\text{dye}} \cdot A_{280}} \cdot 100[\%] \quad (4.4)$$

where A_{dye} is the absorption at the absorption maximum of the dye, $\epsilon_{\text{protein}}$ is the extinction coefficient of the protein ($\epsilon_{\text{myPGK}} = 20.580 \text{ l cm}^{-1} \text{ mol}^{-1}$), ϵ_{dye} is the extinction coefficient of the dye ($\epsilon_{\text{Alexa488}} = 71.000 \text{ l cm}^{-1} \text{ mol}^{-1}$, $\epsilon_{\text{Alexa647}} = 239.000 \text{ l cm}^{-1} \text{ mol}^{-1}$) and A_{280} is the absorption measured at 280nm. The extinction coefficient of the protein can be calculated by the number of tryptophanes (w), tyrosines (y) and cysteines (c) (Gill & Hippel 1989) expressed by equation (4.5).

$$\epsilon_{\text{protein}}(280\text{nm}) = 5690 \cdot w + 1280 \cdot y + 120 \cdot c \quad (4.5)$$

4.1.4.4 Labelling procedure for acceptor/donor – only myPGK

After the labelling procedure for the double labelled myPGK was finished also a donor-only myPGK and acceptor-only myPGK was prepared. Therefore, 4 ml myPGK in labelling buffer were concentrated to 1 ml (23 μM myPGK). The concentration was determined according to the Warburg Christian equation (3.4). Then, the myPGK solution, 30 ml phosphate buffer and 1 ml 1 M NaOH were degassed under vacuum for 20 minutes. In addition, the solutions were bubbled with nitrogen for 15 minutes. 3 mg TECP were dissolved in 800 μl phosphate buffer and 200 μl 1M NaOH. As mentioned before, 3 mg glutathione was stored at 4 °C for the next day. The labelling reactions of

both dyes were started by combining the components of Table 4.4 separately in two eppis. For the donor-only batch Alexa488 was used and for the acceptor only batch Alexa647.

Table 4.4: Composition of the acceptor/ donor – only protein mixtures

component	concentration	note
acceptor- only myPGK/ donor –only myPGK		
myPGK in LB	10 μ M	
TCEP	200 μ M	20 x the protein concentration
Alexa647/ Alexa488	30 μ M	3 x the protein concentration
Phosphate buffer		Up to 1 ml

The headspaces of both eppis were blown with nitrogen before sealing and the mixtures got vortexed for one minute. The reactions were performed in the dark at room temperature for 12 hours.

On the next day, the prepared glutathione was solved in 1ml distilled water and each labelling reaction was stopped by 300 μ M glutathione. In the next step, both PD-10 columns were equilibrated with phosphate buffer. Each reaction mixture was applied to a column to remove the free dye. At the end, twenty 0,5 ml fractions were collected. For the donor-only myPGK the absorptions at 260 nm, 280 nm and 495 nm were measured and the fractions 3-7 were combined. For the acceptor-only mypgk the absorptions 260 nm, 280 nm and 650 nm were determined and the fractions 2-6 were combined. Then, the PD-10 columns were equilibrated with MOPS buffer and the samples were applied to them. Twenty 0,5 ml fractions were collected again for each column and the absorptions were measured like before. For the donor only myPGK the fractions 3-5 were combined and for the acceptor only myPGK fraction 2-4.

At this point, it was assumed that single or double labelled proteins will give equal fluorescence results. Thus, further cleaning and separation procedures were skipped. Only the condition of covalently bound dyes was relevant without the presence of free dyes in the solution. The concentrations were calculated by the measured absorptions of the dyes which was 3,96 μ M (WC) for the donor only myPGK and 3,71 μ M (WC) for the acceptor only myPGK. Consequently, it was not possible to distinguish whether the dye was bound to position 135, 290 or both positions. The labelling efficiency of the donor only solution was 32,3 % and for the acceptor only solution it was 48,1 %.

4.1.5 Structural and functional validation (CD, enzyme activity, DSC)

The next step of the workflow includes the structural and functional validation of the purified and expressed protein. Therefore, circular dichroism was used to compare the composition of secondary structures of the expressed protein with the literature (Rosenkranz 2011a). Furthermore, the structural changes caused by thermal denaturation were monitored by CD. In the next step, the functional activity of the expressed (myPGK) and labelled protein (myPGKdl) was verified in comparison to the wildtype (wtPGK) by an enzyme activity test. The positive results of this test proof that the mutations and the labelling procedure do not have a relevant impact on the native structure of the protein. At last, information about the structural stability as well as thermodynamic parameters of the unfolding pathway were acquired with the DSC. For all these methods the modified yeast phosphoglycerate kinase (myPGK) in the folding studies buffer was used to make the structural and functional validation. Further details are described in the following section.

4.1.5.1 Circular dichroism (CD)

CD was used to verify the structure of the expressed protein and to monitor the changes in absorption of secondary structures during thermal denaturation. Therefore, a CD spectrum of myPGK in FB (8,07 μM) was recorded at 25 °C to verify the successful expression and purification of myPGK. However, the first measurement was carried out after the xenon lamp reached the operating temperature (30min). Furthermore, the CD instrument was flooded with nitrogen for at least 15 minutes to remove the air. The cuvette with the sample was sealed with Parafilm[®] to avoid evaporation of the liquid. The recorded data were analysed according to the paper of Greenfield (2006). Thus, every result with a high tension over 1000 was discarded. Then, the baseline (FB) was subtracted from the sample data. In addition, the sample data was corrected so that the average of the data points between 250 nm and 260 nm was equal to zero. Finally, the whole data set were converted by equation (3.7) to calculate the mean residue ellipticity. The outcome (Figure 4.7) shows that the expression of myPGK was successful because of the high similarity between the analysed data and the literature.

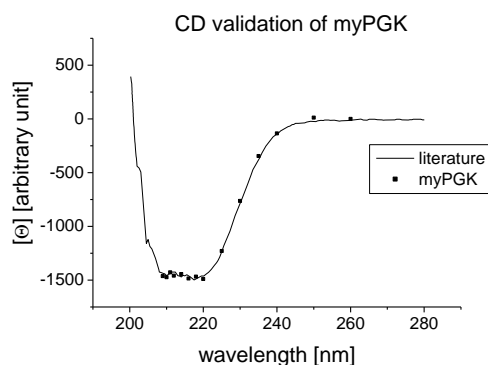


Figure 4.7: Validation of myPGK protein structure, literature: yPGK C97S Q135C S290C (Rosenkranz 2011a)

After the verification, CD spectra at different temperatures and selected wavelengths were recorded. The measurement of a CD spectrum was started after the desired temperature has been reached and took at least five minutes. Consequently, the thermal denaturation was monitored within the temperature range from 35 °C to 71 °C in 3 °C steps. The results are summarised in Figure 4.8.

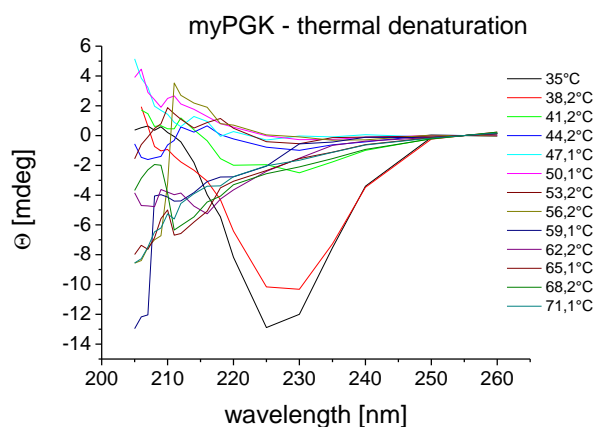


Figure 4.8: CD spectra of the thermal denaturation of myPGK

As a consequence, the temperature dependency of the wavelength area between 218 – 240 nm were analysed for a better interpretation (Figure 2.1).

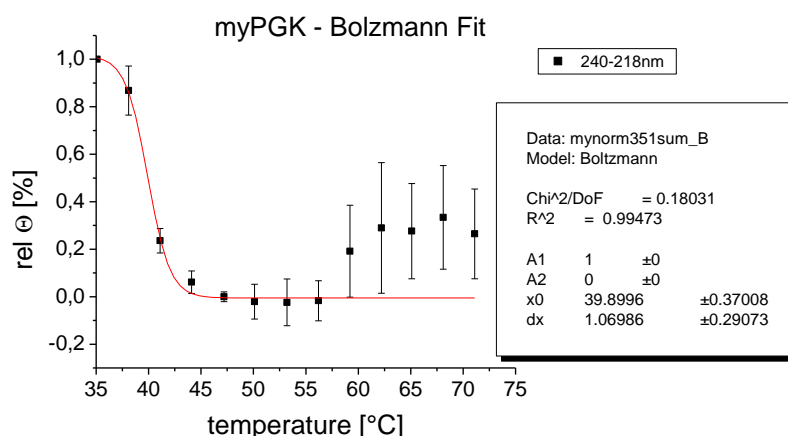


Figure 4.9: CD - temperature dependent absorption of myPGK within 218-240 nm

This wavelength range is relevant because the signals of the secondary structures (α - helices and β -sheets) are interfering constructively. A Boltzmann Fit of the data revealed a melting temperature at 39,9 °C. This means that half of all secondary structures within the protein ensemble were denatured and did not have a contribution to the signal. Furthermore, all structures in the protein disengaged over 47 °C. Other structures formed over 56 °C assuming that effects of aggregation were initiated (Galisteo et al. 1991).

4.1.5.2 Enzyme activity

The functional evaluation of the expressed and purified protein was verified by measuring the absorption of NADH (340 nm) within a coupled optical test as a function of time (p. 30). Therefore, the reaction velocities of the expressed (myPGK) and labelled protein (myPGKdl) were compared to the enzyme activity of the wild type phosphoglycerate kinase of *Saccharomyces cerevisiae* (wtPGK).

The composition of the used coupled optical test (Table 3.1) was chosen to achieve a measurable decrease in absorption of $\Delta\text{OD} \sim 0,3 \text{ min}^{-1}$. Furthermore, the concentration of NADH was 0,2 mM and the other components were present in an excess of 5 mM. The GADPH activity was 11 U for every test batch to eliminate the risk that the measured activity for the expressed and labelled proteins could be influenced by the indicator reaction. In addition, the protein concentrations of all samples were 45 μM to

allow a direct comparison of the reaction velocities. The coupled optical tests were performed at room temperature and the results are shown in Figure 4.10.

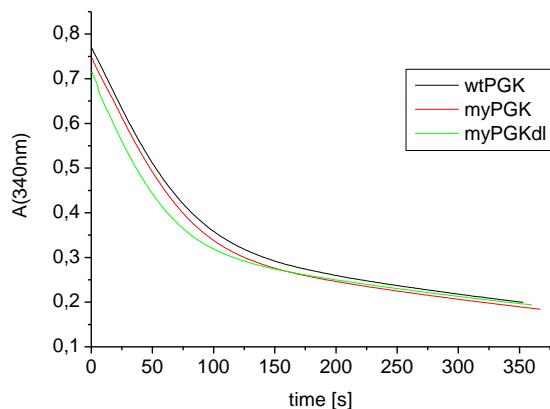


Figure 4.10: results of the coupled optical test

The determination of the enzyme activities were achieved by calculating the reaction velocities between 15 and 30 seconds according to equation (4.6):

$$V = \frac{\Delta OD \cdot V_K}{\varepsilon \cdot d} \cdot 10^3 \quad (4.6)$$

where V is the reaction velocity in U ($=\mu\text{mol}/\text{min}$), ΔOD is the absorption difference in min^{-1} , V_K is the solution volume in the cuvette, ε is the molar-extinction coefficient of NADH at 340 nm which is $6292 \text{ l mol}^{-1} \text{ cm}^{-1}$ (Ziegenhorn, Senn & Bücher 1976) and d is the thickness of the cuvette (1cm).

The resulting reaction velocities are 52 mU for wtPGK, 51 mU for myPGK and 53 mU for the labelled myPGK. This outcome also indicates that the expression, purification and labelling procedure of myPGK was successful to gain functional proteins with a native structure.

4.1.5.3 Differential Scanning Calorimetry (DSC)

The DSC was used to validate the stability of myPGK and to determine thermodynamic parameters. The measurements were performed according to the DSC user manual (MicroCal n.d.) and involved the following steps: *cleaning, sample preparation, the filling procedure, measurement and refilling.*

First of all, the used needles and magnets were cleaned twice by methanol, chloroform, methanol and distilled water. Finally, they were equilibrated in FB buffer. The DSC cells were cleaned at 50 °C with 50 ml water, flushed with air, cleaned with 50 ml detergence (0,5 %), then again with 50 ml water and dried with air. Before every measurement, the temperature of the cells was adjusted to 25 °C. In addition, the sample and the reference buffer were prepared by degassing and stirring for 15 minutes.

Then, the most critical step was the filling procedure. Therefore, a filling funnel was used to achieve a good repeatability. This part allows the filling syringe needle to stop 1-2 mm above the cell bottom. At the beginning, the syringe was filled very slowly by avoiding any bubbles with the prepared sample solution. Then, the filling syringe needle was inserted through the filling funnel into the DSC cell. The syringe plunger was slowly depressed with a constant pressure until 0,5 ml were injected. A few short bursts were performed to push out any remaining bubbles from the working volume of the cell. In the next step, the needle and the filling funnel were pulled out avoiding any vibrations. The same procedure was repeated to fill the reference cell with the folding studies buffer. At the end, the excess solutions were removed with the long needle syringe to achieve the same filling volumes in the cells. After the filling, the DSC was sealed with the cap and the o-ring sealed piston was turned clockwise until the pressure was 27 psi.

At this point, the measurements were programmed to run with an upscan rate of 90 °C h⁻¹ from 25 – 63 °C and a downscan rate of 60 °C h⁻¹ from 63 – 25 °C. Consequently, reference data was acquired by measuring FB buffer versus FB buffer. The sample DSC data was recorded by scanning myPGK in FB versus FB. After the first scan was carried out with the protein, the following DSC curves did not show any transitions (Figure 4.11). As a result, the unfolding process was considered to be irreversible at higher temperatures than the transition temperature T_m . This effect of irreversible thermal denaturation was also mentioned for the wildtype (yPGK) by the work of Galisteo et al. (1991) as well as by the work of Strucksberg et al. (2007).

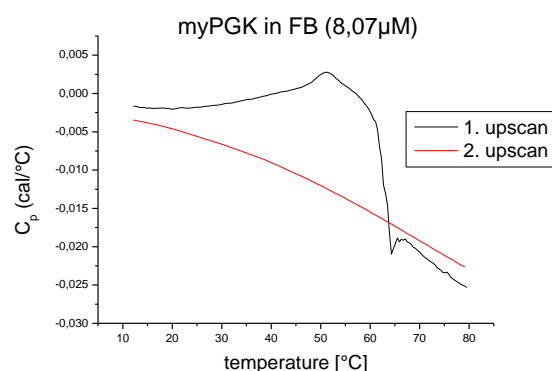


Figure 4.11: Irreversible protein unfolding as a result of protein aggregation

Therefore, the sample solution was diluted by ten times ($0,807 \mu\text{M}$) and it was refilled after every downscan. The first scan was discarded to guarantee the same thermal history for all measurements. Several scans were made until three had similar thermodynamically parameters. An upscan with myPGK in FB is shown in Figure 4.12 A.

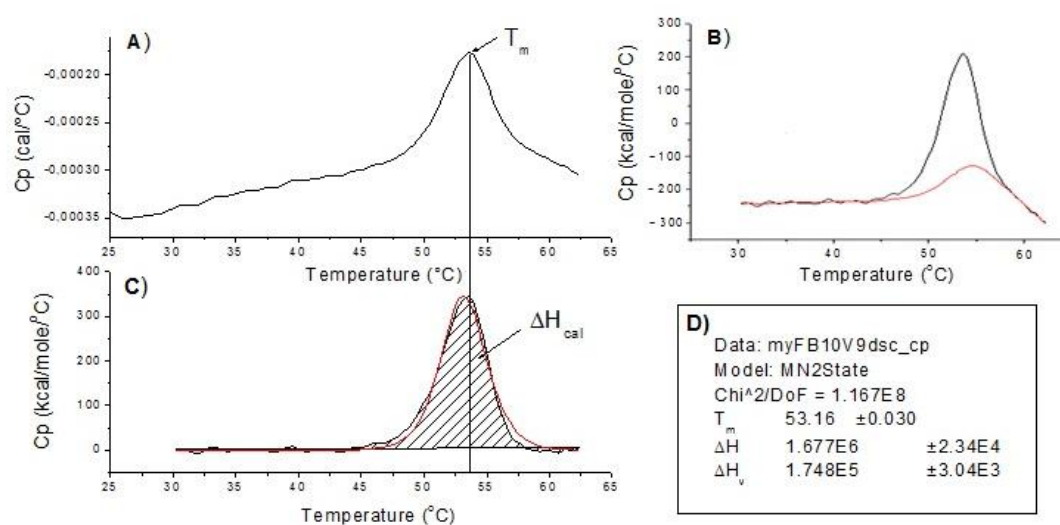


Figure 4.12: Analysis of an DSC upscan with myPGK in FB ($0,807 \mu\text{M}$)
 A) raw data of an DSC upscan; B) Creation of a Progress Baseline after subtraction of the reference data and normalisation, C) Fitting of the curve after the subtraction of the baseline, D) thermodynamic results of the curve fitting, T_m : transition temperature, $\Delta H = \Delta H_{cal}$: calorimetric enthalpy, ΔH_v : van't Hoff enthalpy

The data analysis of the DSC results was performed according to the tutorial guide of MicroCal (1998). This included the deletion of data below 30 °C. Furthermore, the reference data was subtracted from the sample data (Figure 4.12 A) and normalised by the concentration (0,807 μM). In the next step, a chemical baseline (Progress baseline) was created (Figure 4.12 B) and subtracted. At last, a curve was fitted using a non-2-state model (Cursor Init). The thermodynamically result of the curve fits is listed in Table 4.5.

Table 4.5: DSC results of myPGK in FB

T_m [°C]	ΔH_{cal} [kcal mol ⁻¹]	ΔH_v [kcal mol ⁻¹]
$53,16 \pm 0,038$	$1700 \pm 48,06$	$184,0 \pm 14,86$

The results show that $\Delta H_{\text{cal}} > \Delta H_v$. Consequently, the unfolding pathway is more complex and follows not a two-state model ($N \rightarrow U$). It is also important to know that the transition temperature is scan rate dependent (Galisteo et al. 1991). Therefore, the T_m of 53,16 °C is the result of an 90 °C/h upscan rate. Further thermodynamic analysis of the unfolding process is not reliable because a consistent model is not available. Current models based on the results of the literature are mentioned in the discussion chapter (p.101).

4.1.6 Fluorescence and absorption measurements

After the structural and functional validation of the labelled proteins it was possible to determine their spectral properties. Therefore, the next step of the workflow included the fluorescence measurements of the double labelled myPGK (1,65 μM) in folding studies buffer (FB), the donor-only myPGK (0,54 μM) in FB and the quantum yield standard fluorescein (0,48 μM) in 0,1 M NaOH. Besides, absorption spectra were recorded of donor-only myPGK (0,54 μM) in FB, acceptor-only myPGK (1,25 μM) in FB, fluorescein (0,48 μM) in 0,1 M NaOH and myPGK (0,54 μM) in FB for baseline correction. Higher concentrations were generated with a 10.000 MWCO concentrator which was necessary to achieve sufficient fluorescence intensities. However, myPGK concentrations over 3 μM indicated effects of aggregations and were avoided. All measurements were carried out at the following temperature points: 15 °C, 16,5 °C, 18 °C, 21,5 °C, 25,5 °C, 28 °C, 31 °C, 35 °C, 38 °C, 42 °C, 46 °C, 48 °C, 51 °C, 54 °C,

59 °C, 63 °C, 67 °C and 75 °C. Therefore, the fluorimeter or the absorption spectrometer was serial linked to a water bath and a refractometer. This setup made it possible to measure the temperature T , the refraction index n and the fluorescence or the absorption of the sample simultaneously. Practically, it was necessary to place silica gel in the measurement chamber of the fluorescence spectrometer to suppress condensation on the surface of the cuvette. In addition, the cuvette was sealed with Parafilm[®] to avoid evaporation of the liquid. Otherwise, it would have led to a higher concentration and aggregation of the protein. The parameters used for the measurements are summarised in Table 4.6 and Table 4.7.

Table 4.6: Parameters for the fluorescence measurements

sample	excitation wavelength [nm]	emission range [nm]	scan rate [nm/min]	excitation slit width [nm]	emission slit width [nm]
myPGKdl	495	480 - 750	300	5	5
myPGK donor only	495	480 - 750	300	5	5
fluorescein	492	470 - 600	300	5	5

Table 4.7: Parameters for the absorption measurements:

sample	path length [mm]	absorption range [nm]	scan rate [nm/min]	slit width [nm]	interval [nm]
myPGK donor only	10	600 - 420	300	5	1
myPGK acceptor only	10	750 - 480	300	5	1
fluorescein	10	550 - 370	300	5	1
myPGK	10	750 - 420	300	5	1

The results of the fluorescence and absorption measurements are depicted in Figure 4.13.

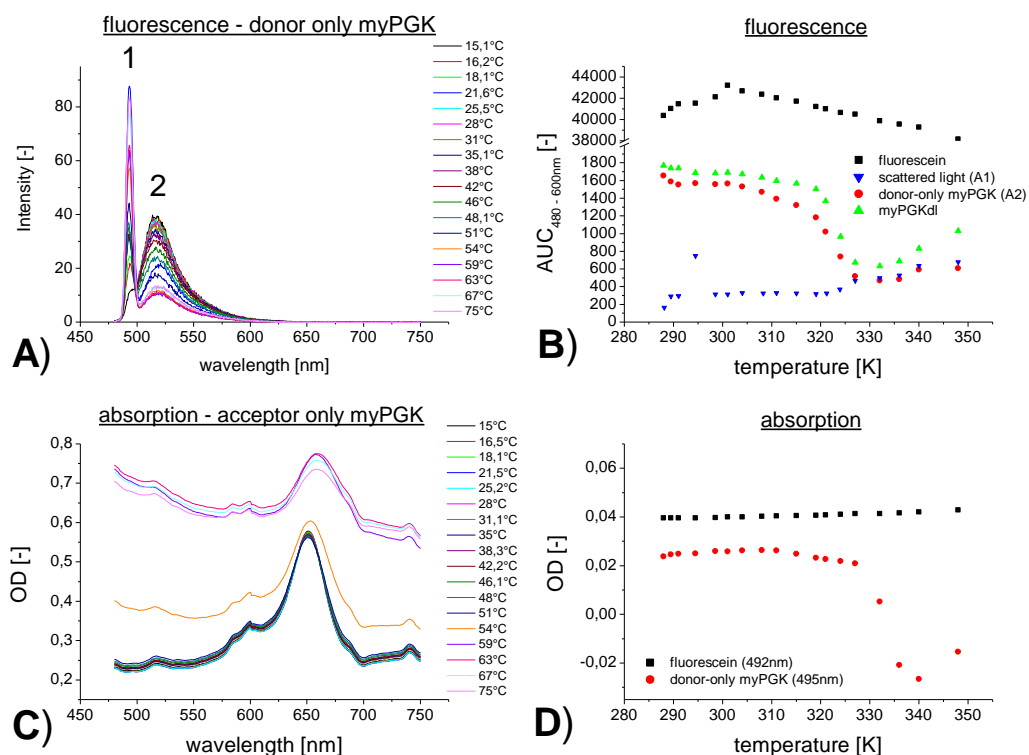


Figure 4.13: Results of the fluorescence and absorption measurements, A) temperature dependent fluorescence of donor only myPGK, 1) scattered light at the excitation wavelength, 2) fluorescence of Alexa488, B) temperature dependence of the integrated fluorescence spectra between 480 and 600 nm, C) absorption of myPGK acceptor only, D) absorption at the excitation wavelengths

The fluorescence spectra in Figure 4.13-A reveals that besides the fluorescence of the donor (peak 2) also the scattered light at the excitation wavelength (peak 1) was recorded. Consequently, the peaks were fitted individually to gain the area under curve (AUC). The outcome of these calculations is summarized in diagram B. It shows that the fluorescence of the chemical compound fluorescein reaches its maximum at 301 K before it drops constantly with the temperature. Apart from that, the fluorescence of the donor attached to myPGK indicates also a slight decrease before it falls more (double labelled) or less (donor only) rapidly after 319 K. At temperatures over 327 K it appears that the fluorescence of the donor has the same progress as the scattered light. According to the DSC results it corresponds to the properties of aggregation which falsify the data. Furthermore, this effect is also visible in the absorption spectra of the

donor and is illustrated as significant change in diagram C and D after 327 K. Hence, the absorption of fluorescein is minimal affected by the temperature. At last, the results of the refractive index measurements are depicted in Figure 4.14. This diagram demonstrates that the temperature has a minor effect on the refractive index of the folding studies buffer and the 0,1 M NaOH.

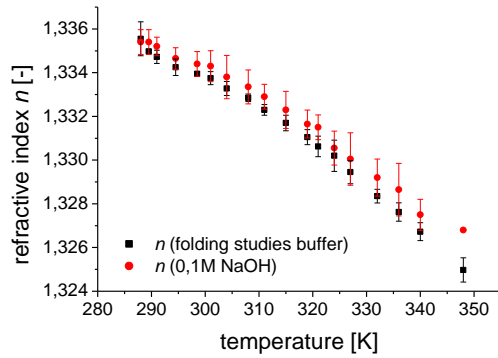


Figure 4.14: Temperature dependence of FB and 0,1 M NaOH reference indices

These spectral properties were used to calculate the quantum yield of the donor Q_D and the spectral overlap $J(\lambda)$. Moreover, these values were essential to determine the temperature dependent Förster radius R_0 .

4.1.7 Quantum yield and spectral overlap

For this work, the quantum yield of the donor Q_D was determined by comparing the spectral properties of donor-only myPGK with those of the standard fluorescein. Therefore, the absorption spectra, fluorescence spectra and refractive indices of the donor-only myPGK and fluorescein were utilized to calculate the quantum yield at different temperatures according to equation (4.7):

$$Q_D = Q_R \cdot \frac{I}{I_R} \cdot \frac{OD_R}{OD} \cdot \frac{n^2}{n_R^2} \quad (4.7)$$

where Q_D and Q_R are the quantum yields of Alexa488 and fluorescein ($0,95 \pm 0,03$) in 0,1 mol/l NaOH (Brannon & Magde 1978). I and I_R represent the integrated fluorescence spectra of donor-only myPGK and fluorescein (Figure 4.13 B). OD_R and OD are the absorption values at the respective fluorescence excitation wavelengths (Figure 4.13 D). n and n_R are the refractive indices of the folding studies buffer and the 0,1 M NaOH (Figure 4.14). Besides the quantum yield, the spectral properties were also

utilized to determine the spectral overlap $J(\lambda)$ between the labelled dyes Alexa488 and Alexa647. Therefore, the fluorescence spectra of donor only myPGK (Figure 4.13-A, peak 2) and the absorption spectra of acceptor only myPGK (Figure 4.13-C) were used to calculate $J(\lambda)$ between 525 and 635 nm according to equation (3.15). The temperature dependent results of Q_D and $J(\lambda)$ are illustrated in Figure 4.15.

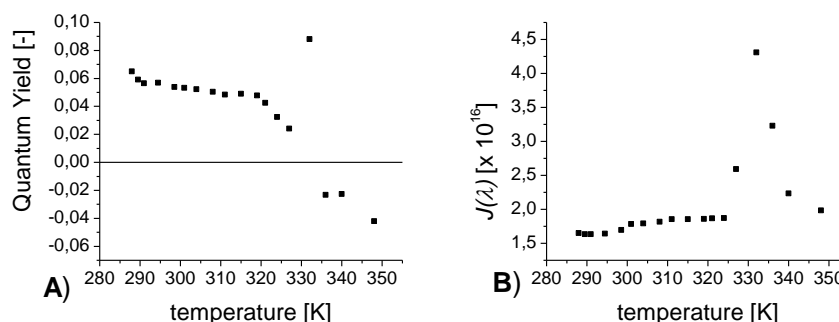


Figure 4.15: Temperature dependency of the donor's quantum yield Q_D (A) and the spectral overlap $J(\lambda)$ (B)

As a consequence, graph A illustrates that the protein has a major influence on the quantum yield of the donor. Usually, Alexa488 has a quantum yield of 0,92 (Life Technology Corporation 2013b). Nevertheless, it appears that the nearby protein reduces it to 0,06. In addition, the temperature dependent aggregation of the protein reduces the quantum yield even more after 319 K and lead to unreliable results after 327 K. This influence is also visible in graph B and results in significant changes after 324 K.

4.1.8 Efficiency of energy transfer – EET

4.1.8.1 EET from fluorescence intensities

The results of the fluorescence measurements of myPGKdl and donor-only myPGK were utilized to calculate the EET_{flu} according to equation (3.18). Therefore, the fluorescence intensities of myPGKdl and donor-only myPGK were converted to the same concentration. A correlation between the measurement points was established by fitting the result with a Boltzmann function (Figure 4.16).

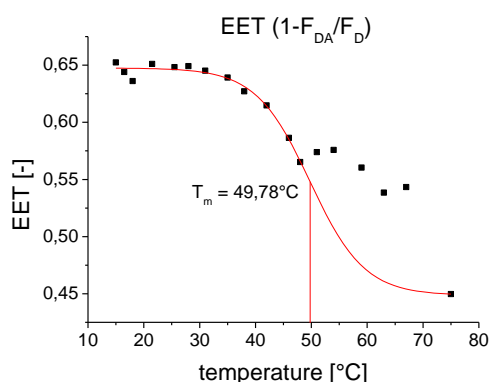


Figure 4.16: EET resulting from fluorescence measurements

The fit indicates that the melting temperature T_m is at 49,78 °C and the transition changes with a steepness of 9,75 mEET/°C. Hence, for this function it was assumed that the data point at 75 °C represents the measured EET minimum. Unfortunately, the outcome of previous measurements reveals that the results over 54 °C are not reliable because the protein starts to aggregate. Therefore, the melting temperature T_m of the EET_{flu} data points is an extrapolated value. Nevertheless, the results also show that the interdomain distance increases until 48 °C before it decreases a little. Finally, the distance between the domains increases again after 53,5 °C.

4.1.8.2 EET from fluorescence lifetime

In cooperation with the Hungarian Academy of Science it was possible to determine the fluorescence lifetime of donor-only myPGK τ_D and myPGKdl τ_{DA} . The results were utilized to calculate a more reliable EET according to equation (3.17). Practically, the first step of this procedure included the determination of the instrument response function (IRF). The IRF was used to correct the data for internal effects of the fluorimeter. Therefore, the calibration was performed with a beam stopper in the instrumental setup and Ludox[®] (30 % silica gel) as sample. Thereupon, the decay measurements of myPGKdl (0,925 μ M) and donor-only myPGK (2,97 μ M) were executed with a polariser in the emission beam at 54,8° and a 495nm cut-off filter. The labelled proteins were excited with the EPL-470 Picosecond Pulsed laser at 472 nm. The fluorescence was recorded with 2048 channels at a 5 nm slit width and the detection was stopped after 10000 counts. Previous methods lead to the assumption that the protein starts to aggregate over 54 °C. Therefore, all measurements between 15 °C and 51 °C were performed at first before records were made over 54 °C. The resulting

decays were analysed according to a biexponential function and the average lifetimes were calculated according to equation (3.12). Furthermore, the average lifetimes were used to calculate the EET. The analysed results are illustrated in Figure 4.17-A.

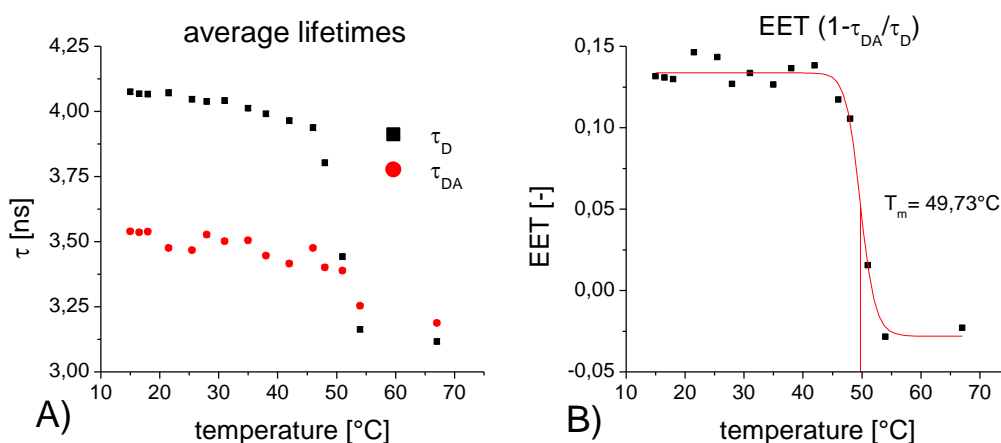


Figure 4.17: Fluorescence lifetime measurements, A) temperature dependent lifetimes of myPGKdl (τ_{DA}) and donor-only myPGK (τ_D), B) temperature dependent efficiency of energy transfer, T_m ... melting temperature

The average lifetime of donor-only myPGK correlates with $\tau_{\text{Alexa488}} = 4.1$ ns of the literature (Life Technology Corporation 2013b). Hence, the resonance energy transfer has a smaller effect on the average lifetime of Alexa488 within myPGKdl as expected. Therefore, the calculated EET is relative small. According to the work of Dr. Rosenkranz (2011a) the EET in the native state of myPGK should be at least between 0.6 and 0.8. This would correspond with a lifetime between 1.64 – 0.82 ns. The average lifetimes were calculated as sum of the partial lifetimes. The conclusion of these results is that the protein solution of myPGKdl also contained a significant amount of free donor. Consequently, the last purification step with the PD-10 desalting column was not efficient enough. Furthermore, for fluorescence lifetime measurements it should be considered to use the DEAE column again after the second labelling procedure. Nevertheless, the calculated EET correlates with a Boltzmann fit as depicted in Figure 4.17-B. This fit reveals that the protein has also a melting temperature T_m at 49.73 °C. At this transition point, the concentrations of folded and unfolded proteins are equal. The Boltzmann fit decreases at a rate of -0.028 EET/°C. Negative EET results over 51 °C underlines the assumption that the protein starts to aggregate.

4.2 Simulation

The general goal of the computer simulations was to acquire information about the temperature dependent distance distribution of the labelled fluorophores Alexa488 and Alexa647 and the relative orientation of their transition dipoles κ^2 . Those dyes were covalently bond to the positions 155 and 310 of the modified yeast phosphoglycerate kinase. The data was generated by molecular dynamic simulations (MD) and the desired values were calculated based on the trajectories.

4.2.9 Homology Modelling

In order to start the simulations of modified yeast phosphoglycerate kinase (myPGK) a homology model was generated with the sequence given by Dr. Rosenkranz (p. 113). For this procedure a sequence similarity search was performed with the online search engine BLAST. As a result, the structures 1QPG, 2K4Z and 1DVJ from the Protein Data Bank (PDB) were chosen as templates. 1QPG has 411 of the 454 amino acids in common with myPGK and represents the main structure. 2K4Z has an identical sequence with the N-terminus and 1DVJ with the C-terminus of myPGK. Subsequently, the program Modeller v9.10 (2012) was used for homology modelling. The procedure was executed according to the tutorial of Sali (2012). Therefore, the next step was the sequence alignment. The result of the alignment is presented on page 113. Consequently, five new models were built after the template was created for myPGK. At last, the models were evaluated according to their Discrete Optimized Protein Energy (DOPE). The DOPE score is a relative value which allows the separation of good and bad models. Models with the lowest DOPE score have the highest confidence. The results indicated that the fifth model was the best according to the global DOPE score listed in Table 4.8.

Table 4.8: DOPE score results of the multiple template models

multi-template model	Global DOPE score
myPGK model 1	-42741
myPGK model 2	-43739
myPGK model 3	-43890
myPGK model 4	-43177
myPGK model 5	-44272

However, the comparison of the DOPE score profiles reveals deviations between the fifth model and the main template 1QPG. The biggest differences are around the alignment positions 200 and 400 as displayed in Figure 4.18.

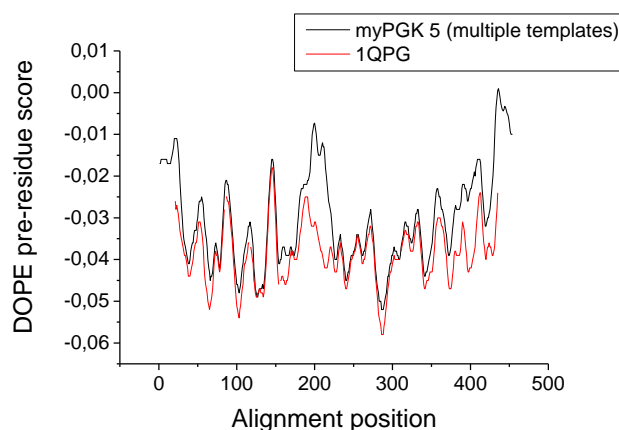


Figure 4.18: Comparison of the DOPE score profiles of the fifth myPGK-model and the main template 1QPG

4.2.10 Molecular dynamic simulation

The simulations of the wild type yeast phosphoglycerate kinase (wtPGK) and the modified yeast phosphoglycerate kinase (myPGK) were performed with the GRONINGEN MACHINE for CHEMICAL SIMULATIONS (GROMACS v.4.5.3) (Spoel et al. 2010) on the VIENNA SCIENTIFIC CLUSTER 1 (VSC – 1). The molecular dynamic simulation of the doubled labelled modified yeast phosphoglycerate kinase (myPGKdl) with Alexa488 at position 155 and Alexa647 at position 310 were executed with the high-speed MD program Desmond (2012) from the Maestro software package.

4.2.10.1 GROMACS (wtPGK, myPGK)

The setup for the MD simulations of wtPGK and myPGK were carried out according to the GROMACS tutorial (Lemkul 2012). The different steps of the setup, the MD run and the analysis are depicted in Figure 4.19.

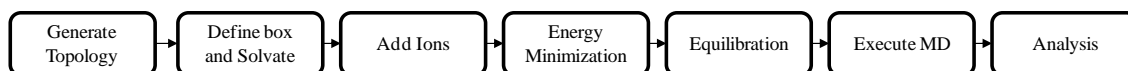


Figure 4.19: Scheme of the molecular dynamics simulation procedure (Lemkul 2012, modified)

The specific parameters used for every setup step are summarised in the appendix (p. 114). At first, the structure of wtPGK was generated based on the Protein Data Bank (PDB) file 1QPG. It represents the R65Q mutant of wtPGK. Consequently, the mutation was reversed with the PDB Swiss Viewer (Guex & Peitsch 1997). Furthermore, the substrates within the 1QPG file were removed with MOE (2008). The structure of myPGK was generated by homology modelling (p. 88).

In the next step, the topology file was created using the OPLS-AA force field. The charge states had to be assigned manually for the amino acids glutamic acid (Glu), aspartic acid (Asp), lysine (Lys), arginine (Arg) and histidine (His). Therefore, the structures were tested online with PROPKA 3.0 (Søndergaard et al. 2011) to calculate the pK_a values of the polar amino acid. Furthermore, a pH = 7,4 was defined for the system in accordance with the experiment. The charge states were determined according to the Henderson- Hasselbalch equation (4.8) (Hug & Reiser 2000).

$$\text{pH} = \text{pK}_a + \log\left(\frac{[\text{A}^-]}{[\text{HA}]}\right) \quad (4.8)$$

If the result is pK_a < pH, the residue is deprotonated irrespective of whether the residue is an acid or a base. Otherwise, if the outcome is pK_a > pH, the residue is protonated. Consequently, a deprotonated acid (e.g. ASP, GLU) and a protonated base (e.g. LYS, ARG, HIS) are charged. Furthermore, the hydrogen atoms were ignored in order to build the topology file.

At next, the dimensions of the cubic simulation box were defined so that the protein was at least 1 nm away from the edges. It was filled with water using the three point water model SPC. Also the periodic boundary conditions were used for the simulations. The third step included the addition of ions like sodium and chloride to neutralize the charges of the protein. The fourth step included the energy minimization (EM). The EM was stopped after the maximum force was <1000.0 kJ/mol/nm or 50000 iterations were performed. The output of the wtyPGK energy minimization is shown in Figure 4.20.

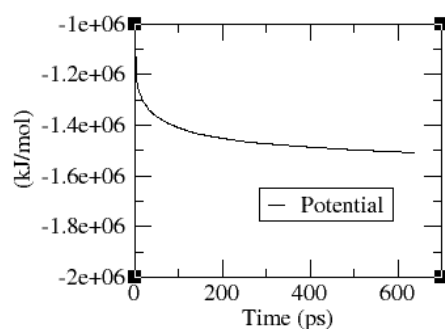


Figure 4.20: Energy minimization of wtyPGK

In the fifth step, the system was equilibrated after the EM. Therefore, position restraints on the heavy atoms were applied to equilibrate the solvent around the protein. The equilibration was performed in two phases. The first equilibration step included the canonical ensemble (NVT) which correlates to isothermal-isochoric conditions. At this point, the desired temperature was adjusted for all simulations with the velocity-rescale algorithm. Simulations for wtyPGK and myPGK were performed at the following temperatures: 298 K, 308 K, 317 K, 326 K, 335 K and 344 K. The NVT step was performed for 100 ps. The result of a NVT equilibration was evaluated according to the output as illustrated in Figure 4.21.

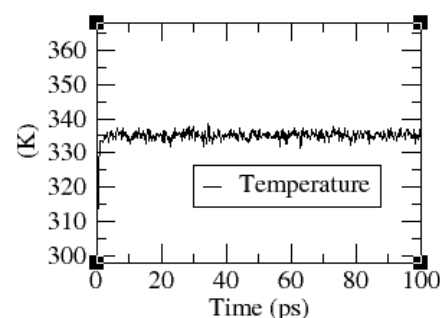


Figure 4.21: NVT equilibration of wtyPGK at 335 K

The outcome shows that the system reached the desired simulation temperature. The second equilibration step included the isothermal-isobaric ensemble (NPT). Therefore,

the system was equilibrated for 100 ps to achieve a constant pressure within the system. For this step, the Parrinello-Rahman barostat was utilized. The results of the NVT equilibration were used for the NPT run. The results are depicted in Figure 4.22.

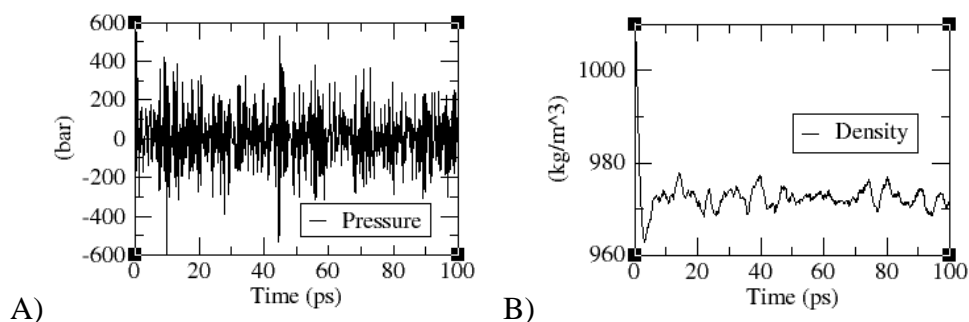


Figure 4.22: A) NPT equilibration of wtyPGK, B) Density of the wtyPGK system

Finally, the system was ready to perform the molecular dynamic simulation. Therefore, any restrains were removed and a simulation of 30ns was performed. The job was executed on the VSC 1 cluster with 256 CPUs in parallel. At the end of the simulation, the results were analysed according to the RMSD distribution of the protein's backbone.

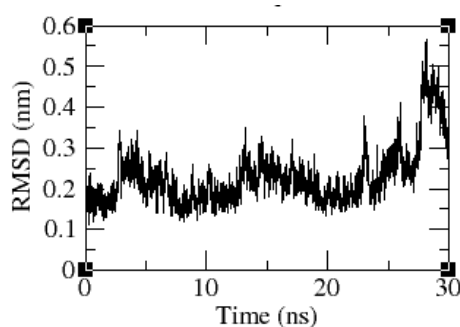


Figure 4.23: RMSD of the backbone during the MD simulation

The results showed that the protein had not reached the equilibrium. Therefore, the simulations were prolonged for another 70 ns. The results are depicted in Figure 4.24.

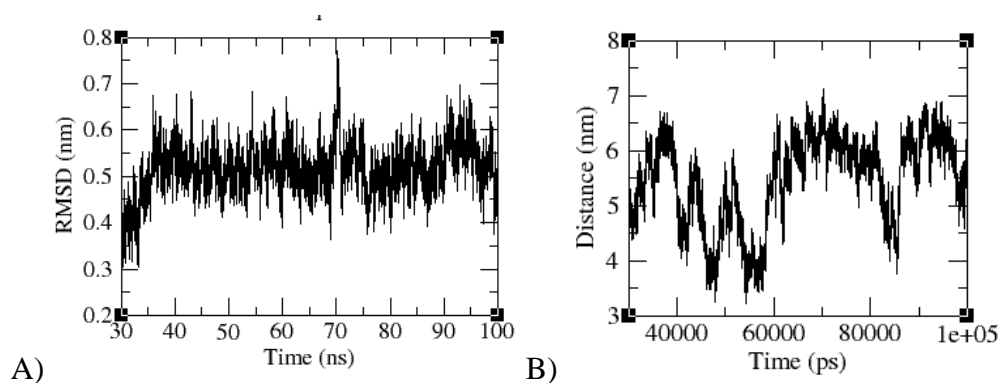


Figure 4.24: MD simulation results of wtPGK from 30 to 100 ns, A) RMSD of the proteins backbone, B) distance between amino acid 135 and 290

The result shows that the system is equilibrated after 35 ns according to the RMSD of the backbone. At this point, the distance between the labelling positions 135 and 290 was calculated. Finally, the temperature dependent distance distribution of myPGK was compared to that of wtPGK. The results should clarify if the mutations have an effect on the movement. The summary of all simulations and their temperature dependent distance distributions is illustrated in Figure 4.25.

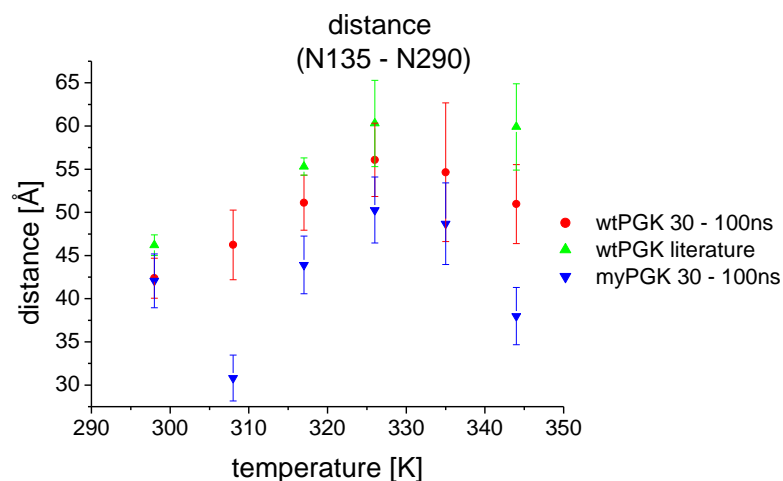


Figure 4.25: Results of the MD simulations, temperature dependent distance distribution of wtPGK, myPGK and the literature (Lillo et al. 1997)

The result indicates that myPGK has a similar progress of the temperature dependent distance distribution between the nitrogen of cysteine 135 and the nitrogen of cysteine 290 as wtPGK. The simulated distances demonstrate also an analog trend in

comparison to the literature and its experimental obtained data (Lillo et al. 1997). The absolute difference between the simulated and measured values is assumed to be the result of different distance measurements. For the simulations the distance between the two labelling positions (nitrogen of amino acid 135 and nitrogen of amino acid 290) were measured. The results of the literature were obtained from FRET measurements using the stopped-flow technique with guanidine hydrochloride. The assumption is that the additional distance corresponds to the size of the linkers between the chromophores of the fluorescence dyes and the protein backbone.

Consequently, the results confirm that the simulations of myPGK are reliable and the mutations have minor effects on the domain movements. The next step included the simulations of the temperature dependent distance distribution of the doubled labelled myPGK. Unfortunately, GROMACS required a manually parametrization of the dye molecules in order to perform the simulations. The parametrization was avoided by using the high speed MD program Desmond.

4.2.10.2 Desmond (myPGKdl)

Desmond (2012) was used to perform MD simulations of the double labelled modified yeast phosphoglycerate kinase (myPGKdl). Those were generated at the same temperature points as for the experiment. The simulations were executed at 288 K, 289 K, 291 K, 294 K, 298 K, 301 K, 304 K, 308 K, 311 K, 315 K, 319 K, 321 K, 324 K, 327 K, 332 K, 336 K, 340 K and 348 K. The MD procedure for myPGKdl was carried out according to the Desmond tutorial (D. E. Shaw Research 2012). The scheme to set up and perform a MD simulation with Desmond is similar to the procedure of GROMACS. The necessary steps are summarized in Figure 4.26.

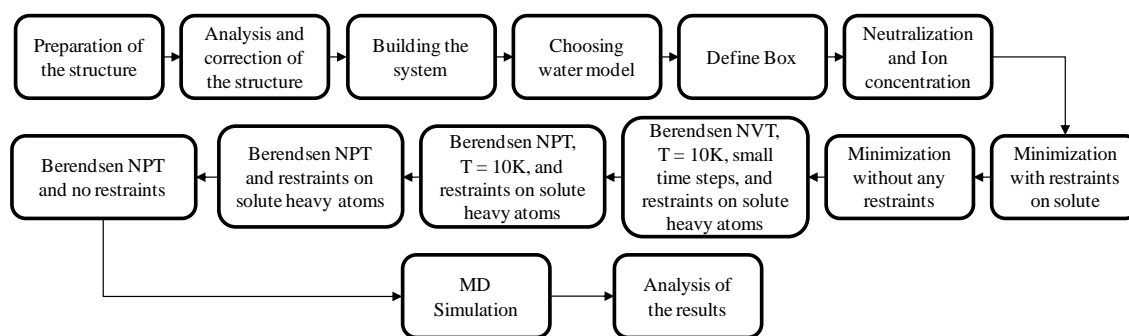


Figure 4.26: Molecular dynamics simulation procedure of Desmond (D. E. Shaw Research 2012)

At first, the structure of the double labelled myPGK was modelled with Maestro (2012). Therefore, the homology model myPGK was complemented by the molecule Alexa488 in position 155 and Alexa647 in position 310 (Figure 3.20). In addition, the structure was analysed for errors like missing atoms or bonds which have been corrected. In the next step, an orthorhombic box was used for the simulations. The edges were defined with a minimum distance of 10 Å to the protein in every direction. Furthermore, the box was filled with the SPC water model. In the next step, ions were added to the system to neutralize the protein's charge. A salt concentration of 50 mM NaCl was chosen to establish experimental conditions. Consequently, the topology file was generated with the OPLS-AA force field. Before any simulations were carried out, the system was minimized and equilibrated according to Figure 4.26 (step 7 – 12). Afterwards, a 37 ns simulation at 298 K was performed to relax the model. This procedure was executed at the VSC 1 with 64 CPUs in parallel. The outcome of this simulation was utilized as starting point for all MD simulations at their respective temperature. Each simulation was generated on the VSC 1 with 256 CPUs in parallel to achieve 100 ns. Unfortunately, the MD runs were terminated after 72 hour and most of them were not finished. Therefore, each simulation was prolonged to reach the 100 ns. Finally, the results were evaluated according to their RMSD values (Figure 4.27).

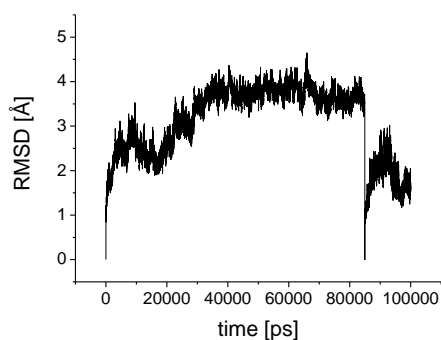


Figure 4.27: RMSD of a 100 ns simulation of myPGKdl

The equilibrium was reached after 30 ns. Therefore, the time frame between 30 and 100 ns of the MD simulations was utilized to obtain the temperature dependent distance distribution for myPGKdl (Figure 4.28). The distances were measured between position 9' of the xanthene of Alexa488 and position 4 of the 1,3 - pentadienyl of Alexa647 (Figure 4.29).

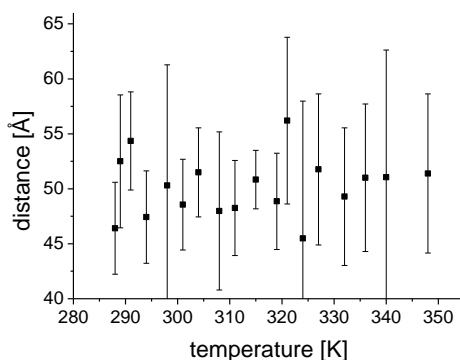


Figure 4.28: Simulated temperature dependent distance distribution of myPGKdl

The results of the simulations show an average distance distribution of 50 Å between the Alexa dyes. Furthermore, the calculations show that the temperature has a minor effect on the interdomain distance. In the next step, the trajectories of the simulations within 30 and 100 ns were used to calculate the orientation factor κ^2 .

4.2.11 Orientation factor – κ^2

The last value to calculate the distance (p. 48) is the orientation factor κ^2 . Usually, it is assumed to be 2/3 for isotropic distributed donors and acceptors and 0,476 for static ones. In order to acquire more realistic FRET results κ^2 was calculated from the trajectories of the MD simulations of myPGKdl.

At first, the dipole moments within the chromophores of Alexa488 and Alexa647 were determined with Maestro (2012). As depicted in Figure 4.29 the dipole within the Alexa488 chromophore points from 9' of the xanthene to the sulphate group in position 5'. The dipole within the Alexa647 chromophore is directed from position 4 of the 1,3 – pentadienyl towards position 3a of the distal indole.

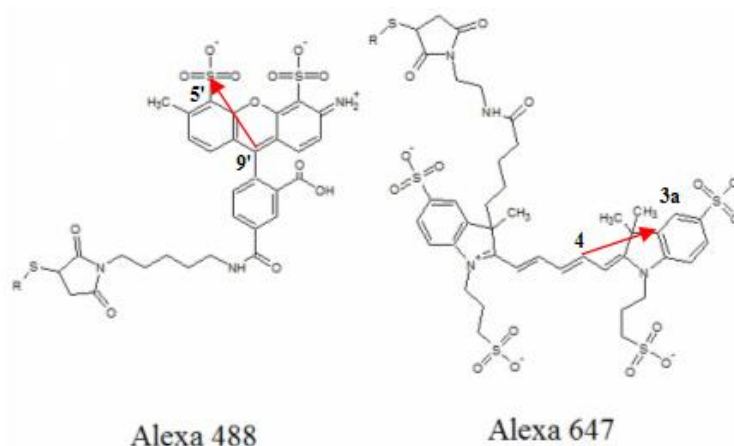


Figure 4.29: Structure of the fluorescence dyes Alexa488 and Alexa647 and the dipole moments of the chromophores

In the next step, the angles θ_T , θ_D and θ_A were determined for every simulation between 30 and 100 ns. Consequently, the orientation factor κ^2 was calculated according to equation (3.35). The results are illustrated in Figure 4.30. The outcome indicates that the temperature has an insignificant influence on the average κ^2 . The orientation factor of the individual temperature point covers a range from 0 to 2,25. Nevertheless, a linear fit of all results reveals that κ^2 is nearly identical with the value assumed for static donors and acceptors ($\kappa^2 = 0,476$). Hence, the whole range of the orientation factor was taken into account in order to calculate the temperature dependent distance r .

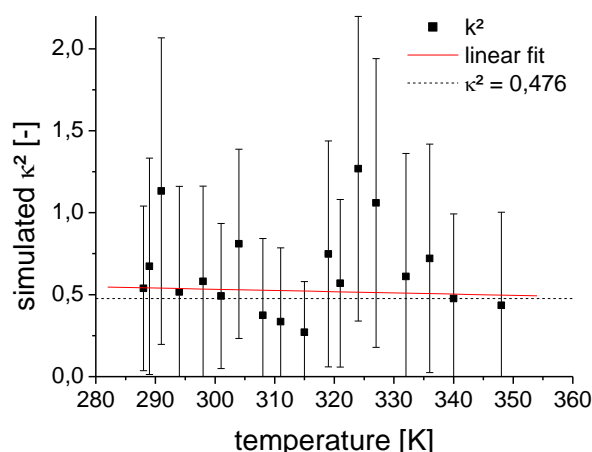


Figure 4.30: Simulated temperature dependency of κ^2 (myPGKdl)

An explanation for the rigid behaviour of the fluorescence dyes was found after the visualisation of the MD simulations. It appears that the sulphate groups of Alexa488 build a salt bridge to the amino group of lysine 168 (Figure 4.31). This ionic bond limits the movement of the chromophore and the resulting dipole. Furthermore, the close proximity of the chromophore to the protein could also be an explanation for the relative small quantum yield (Figure 4.15-A). Hence, the mobility and spectral properties of Alexa647 seems not be affected by the protein.

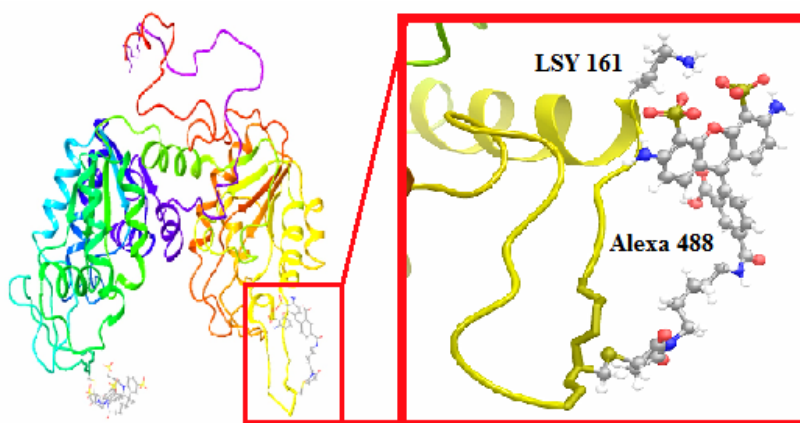


Figure 4.31: Result of the 37 ns relaxation of myPGKdl, left: myPGKdl, right: the negative charged sulphate groups of Alexa488 formed a salt bridge with the positive charged amino group of lysine 161 (Desmond 2012).

4.3 Analysis

4.3.12 Förster radius – R_0

Finally, all required variables were determined in dependence of the temperature to calculate the Förster radius according to equation (3.16). This included the experimental measured refractive index n (Figure 4.14) of the sample solution (FB), the quantum yield of the donor Q_D (Figure 4.15-A), the spectral overlap $J(\lambda)$ (Figure 4.15-B) and the simulation of the orientation factor κ^2 (Figure 4.30). The outcome of the temperature dependent Förster radius is depicted in Figure 4.32.

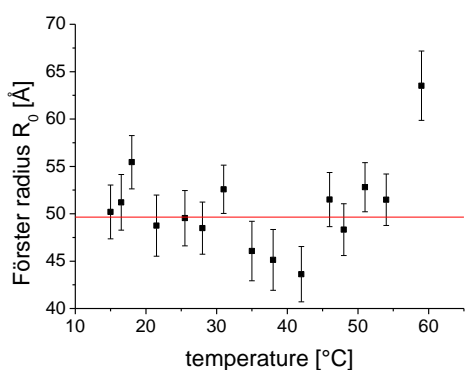


Figure 4.32: Temperature dependent Förster radius R_0

The results show that the Förster radius is randomly distributed between 41 and 67 Å. Significant changes occur over 54 °C which are not reliable because of aggregation. The mean value of 49,65 Å is in accordance with the work of Dr. Rosenkranz (2011a). In his work it was stated that the Förster radius of the free dyes ($R_0 = 53,5$ Å) changes to 49 Å after labelling to the protein.

4.3.13 Temperature dependent distance distribution - r

The last step of this work included the calculation of the temperature dependent distance distribution $r(T)$ between the two domains of myPGK. This variable depends on the Förster radius and the efficiency of energy transfer which are based on the FRET measurements and the simulated results. Therefore, $r(T)$ is calculated according to equation (3.19) and the outcome is represented in Figure 4.33.

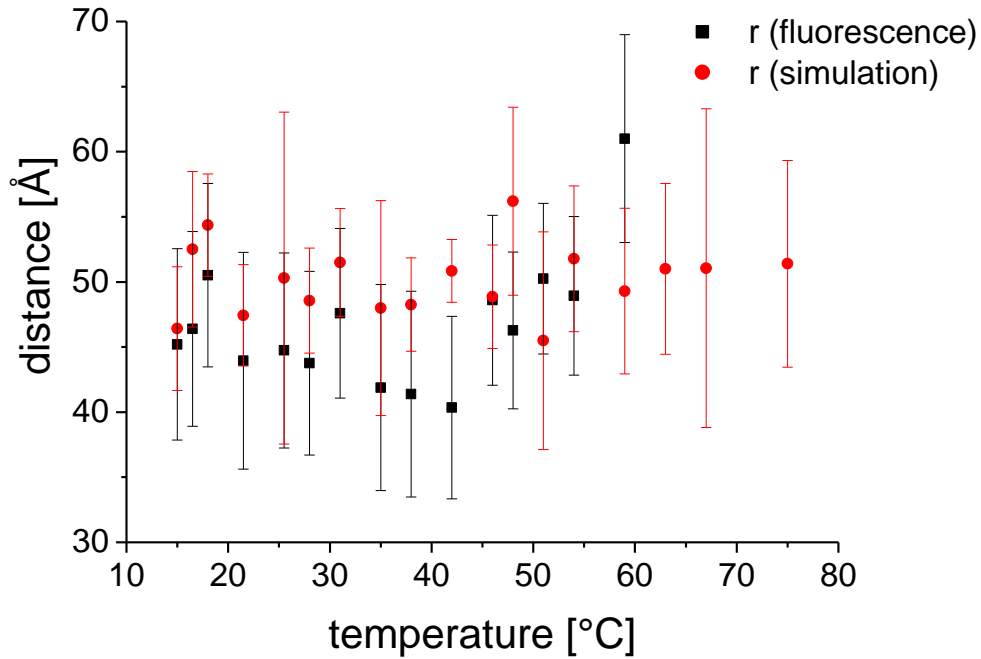


Figure 4.33: Simulated and measured temperature dependent distance distributions between the domains of myPGK

The results show that the interdomain distances of the fluorescence measurements r_{flu} and the simulations r_{sim} are in the same range of 40 to 60 Å. Furthermore, the temperature dependent profile of both methods is similar until 46 °C. At first, r_{flu} decreases with the temperature and reaches a minimum at 42 °C with an average distance of 40 Å. Afterwards, the domains keep an average distance of 48 Å between 46 °C to 54 °C before the interdomain distance increases to 60 Å. Therefore, it appears that the two labelling positions of the protein are about 12 Å further apart after the denaturation.

5 Discussion

The main goal of this work was to acquire information about the temperature dependent unfolding profile of modified yeast phosphoglycerate kinase (myPGK). The unfolding process was characterised within a temperature range of 15 °C to 75 °C by experiments and simulations.

Thus, steady-state CD measurements at various temperature points revealed a melting temperature T_m at 39,9 °C. This point reflects a state where half of the secondary structures vanished within the protein ensemble. Furthermore, over 47 °C the proteins seemed to exist in an intermediate state without any detectable secondary structures. At the end, the CD results confirmed that new structures began to form over 56°C which correlates with protein aggregation. Such an effect was also observed during the DSC upscan measurements and was minimized by diluting the sample solution to a concentration below 1µM. The work of Damaschun et al. (2000) showed that aggregation starts during the transition state with unfolded PGK monomers. The monomers accumulate to octamers and form 73nm long amyloid fibrils with cross- β structures. Besides, the DSC measurements gave a melting temperature T_m at 53,16 °C where half of the proteins were unfolded. Similar results were reported for the wildtype by Galisteo et al. (1991). In this work it was also mentioned that the melting temperature T_m depends on the scan rate and protein concentration. Furthermore, the extracted thermodynamic parameters from the DSC results have also proven that the unfolding process does not follow a simple two-state transition and includes at least one intermediate.

Further information about the thermal unfolding profile was obtained by measuring the interdomain distance via FRET. Therefore, the cysteines 155 and 310 of myPGK were labelled with Alexa488 and Alexa647. In addition, MD simulations of the labelled protein were performed to refine and compare the measured data. As a result, the simulated and measured profiles of the interdomain distance showed a similar progress (Figure 4.33). However, significant deviations were recognisable after the protein starts to unfold and were also limited because of the effects of aggregation. Nevertheless, the profiles revealed that the interdomain distance decreases with the temperature until it

reached an average distance of 40 Å at 42°C. Afterwards, the domains moved apart and remained mainly around 48 Å during the intermediate state between 46 °C and 54 °C. Finally, the results indicated that the proteins started to unfold at higher temperatures before the aggregation was initiated. At the end of this phase, the cysteines moved 12 Å further apart to an average distance of 60 Å. Such an increasing distance between the domains was also reported in the work of Lillo et al. (1997) who used guanidine hydrochloride for chemical denaturation.

6 Conclusion and Outlook

The experimental methods and simulations showed that myPGK unfolds with at least one intermediate due to thermal denaturation. Therefore, the unfolding process can be summarized according to the following scheme: $N \rightarrow I \rightarrow U \rightarrow F$. This folding path represents the transition of the native state N to an intermediate I which is characterised by the lack of secondary structures and an average interdomain distance of 48 Å. In addition, the protein unfolds (U) at higher temperatures, starts to aggregate and remains in an irreversible final state F. Moreover, the literature (Lillo et al. 1997) suggests that the unfolding process is even more complex based on chemical denaturation. Nevertheless, this thesis represents a participation in the project 'Refolding of Biological Macromolecules Initiated by Ultrafast cooling' and the acquired data about myPGK will be utilized as reference. The project analyses the thermal renaturation of proteins by smFRET measurements. Therefore, effects of protein aggregation are circumvented because the sample concentration does not exceed 50 pM. Furthermore, for a detailed analysis of the refolding process of PGK additional interdomain and intradomain distances should be determined by constructing several PGK mutants with different labelling positions. In addition, the method which is used for this project should be validated by investigations on other labelled proteins.

7 Abbreviations & Mathematical Terms

A_L, A_R	absorption of left or right-handed circularly polarized light
CD	circular dichroism
$[\theta]$	mean residue ellipticity [$\text{deg cm}^2 \text{dmol}^{-1}$]
1,3-DPG	1,3- diphosphoglycerate
3PG	DL glyceraldehyde 3 phosphate
a	first half peak width at 10 % of the peak height [ml]
$A_{280}^{1\text{cm}}$	absorption at 280nm and a path length of 1cm
ADP	adenosine diphosphate
APS	ammonium persulfate
A_s	asymmetry factor
ATP	adenosine triphosphate
AUC	area under the curve
b	second half peak width at 10 % of the peak height [ml]
BLAST	Basic Local Alignment Search Tool
BPB	bromphenol blue
c	concentration [mol l^{-1}]
c	speed of light $c = 2,998 \cdot 10^8 \text{ m} \cdot \text{s}^{-1}$
CD	Circular dichroism spectroscopy
CFU	colony forming unit
$C_p(T)$	heat capacity function [$\text{cal mol}^{-1} \text{ }^\circ\text{C}^{-1}$]
$C_{p,N}(T)$	heat capacity of a protein in the native state
$C_{p,pr}(T)$	partial heat capacity
$C_{p,U}(T)$	heat capacity of a protein in the unfolded state
$C_p^{prg}(T)$	progress heat capacity
$\langle C_p(T) \rangle^{exc}$	excess heat capacity
CPU	central processing unit
CV	column volume
d	path length [cm]
DEAE	diethylaminoethyl
dH_2O	distilled water
DHPH	Dynamic high pressure homogenizer
DNA	desoxyribonucleic acid
DOPE	Discrete Optimized Protein Energy
DSC	differential scanning calorimetry
DTT	DL-dithiothreitol
E	energy [J]
<i>E. coli</i>	<i>Escherichia coli</i>

E1, E2	electronic vectors
EDTA	ethylenediaminetetraacetic acid
EET	efficiency of energy transfer
EET _{flu}	efficiency of energy transfer based on fluorescence intensity measurements
EET _τ	efficiency of energy transfer based on fluorescence lifetime measurements
EM	energy minimization
EPR	electron paramagnetic resonance
F	force [N]
F	final state
FB	folding studies buffer
F _D	fluorescence intensity of the donor in absence of the acceptor
F _{DA}	fluorescence intensity of the donor in presents of the acceptor
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
GADPH	glyceraldehyde-3-phosphate dehydrogenase
GROMACS	GRoningen Machine for Chemical Simulations
h	Planck constant $h = 6,63 \cdot 10^{-34}$ Js
HETP	height equivalent to a theoretical plate
HPLC	high performance liquid chromatography
I ₀	fluorescence intensity at time 0
IEX	Ionexchange
IMAC	immobilized metal chelate affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
IRF	instrument response function
J(λ)	spectral overlap [$M^{-1} cm^{-1} nm^4$]
k	rate constant [s^{-1}]
K(T)	equilibrium constant
kat	Katal
kb	kilobase, Unit to measure the size of nucleotides
k _{nr}	sum of all non-radiative decays
k _T	transfer rate of the energy [ns^{-1}]
L	bed height [cm]
LB	labelling buffer
LiNbO ₃	lithium niobate
LINCS	linear constraints solver
m	mass
MC-PMT	microchannel plate photomultiplier tube
MD	Molecular dynamics simulation
MgSO ₄	magnesium sulphate
MOE	Molecular Operating Environment
MOPS	3-(N-morpholino)propanesulfonic acid

MRW	Mean residue weight [g mol ⁻¹]
MWCO	molecular weight cut-off
myPGK	Modified yeast phosphoglycerate kinase (PGK C97S Q135C S290C)
myPGKdl	double labelled modified phosphoglycerate kinase
N _A	Avogadro's constant [6,022141 • 10 ²³ particles mol ⁻¹]
n	refractive index
N _{th}	number of theoretical plates
N	native state
NaOH	sodium hydroxide
NMR	nuclear magnetic resonance
NPT	isothermal-isobaric ensemble with a constant number of particles, pressure and temperature
NTA	nitrilotriacetic acid
NVT	canonical ensemble with a constant numbers of particles, volume and temperature
OD	optical density
OPLS	optimized potentials for liquid simulations
PAGE	polyacrylamide gel electrophoresis
PBC	periodic boundary condition
PDB	Protein Data Bank
PGK	phosphoglycerate kinase
P _i	phosphate
PME	particle-mesh Ewald
PMSF	phenylmethanesulfonyl fluoride
Q	quantum yield
Q _D	quantum yield of the donor without the present of the acceptor
Q _R	quantum yield of a reference substance
r	distance [Å]
R	universal gas constant (= 8,314 J mol ⁻¹ K ⁻¹)
R ₀	Förster radius [Å]
r _{exp} (T)	experimental determined interdomain distance
R _f	relative mobility
RMSD	root mean square deviation
RNA	ribonucleic acid
r _{sim} (T)	simulated interdomain distance
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
S ₀ , S ₁ , ...	energetic state 1, 2, ...
SDS	sodium dodecyl sulfate
SI	International System of Units
smFRET	single molecule Förster resonance energy transfer
SOP	standard operation procedure
t	time
T	temperature [K]

TCEP	tris(2- carboxyethyl) phosphinehydrochloride
TCSPC	time correlated single photon counting
TEMED	N,N,N',N'-Tetramethylethylenediamin
T_m	melting temperature
U	unit, enzyme activity equal to $\mu\text{mol min}^{-1}$
U	unfolded state
UV	ultraviolet
V	Volt
V	reaction velocity
$V(r)$	potential energy function
VIS	visible spectrum
V_R	retention volume
VSC -1	Vienna Scientific Cluster 1
WC	protein concentration according to Warburg and Christian
w_h	peak width at half peak height [ml]
wtyPGK	wildtype yeast phosphoglycerate kinase
α	fraction of the partial lifetime
β	protein concentration [mg ml^{-1}]
$\beta\text{-NADH}$	beta Nicotinamide adenine dinucleotide
Γ	emission rate of the fluorophore
ΔA	Difference in Absorption
ΔG	Gibbs free energy change
$\Delta H_{\text{cal}}(T_m)$	calorimetric change of enthalpy
ΔH_{vH}	van't Hoff enthalpy
$\Delta S(T)$	change of entropy [$\text{J mol}^{-1} \text{K}^{-1}$]
$\Delta\epsilon$	differential molar circular dichroic extinction coefficient
$\Delta\phi$	phase shift
ϵ	molar extinction coefficient [$\text{l mol}^{-1} \text{cm}^{-1}$]
ϵ_A	extinction coefficient of the acceptor [$\text{M}^{-1} \text{cm}^{-1}$]
θ	Ellipticity [$\text{l mol}^{-1} \text{cm}^{-1}$]
κ^2	orientation factor
λ	wavelength [nm]
τ	fluorescence lifetime
τ_D	lifetime of the fluorescence donor [ns]

8 References

- Avestin n.d., *EmulsiFlex-C5*, viewed 06 March 2013, <http://www.avestin.com/English/c5page.html>
- Avestin Inc n.d., *EmulsiFlex: High Pressure Homogenizers*, viewed 06 March 2013, <http://img02.b2b.hc360.com/pic-2/handbook-pic-8/2-8-756328.pdf>
- Bayrhuber, H, Kull, U, Bäßler, U & Danzer, A 1996, '*Linder Biologie*', vol. 1, Gustav Swoboda & Bruder, Wien
- Bisen, PS & Sharma, A 2013, '*Introduction to Instrumentation in Life Sciences*', CRC Press, Boca Raton
- Buchner, J & Kiefhaber, T 2005, *Protein Folding Handbook*, vol. 1, Wiley-VCH, Weinheim
- Brannon, JH & Magde, D 1978, '*Absolute Quantum Yield Determination by Thermal Blooming. Fluorescein*', *The Journal of Physical Chemistry*, vol. 82, no. 6, pp. 705-709.
- California Lutheran University (CLU) 2012, *Protein Analysis-Determination of Protein Concentration*, 14 September, viewed 24 January 2013, <http://public.callutheran.edu/~revie/biochemistry/Protein-analysis-lab.pdf>
- Cutler, P 2004, '*Protein Purification Protocols*', 2nd edn., Humana Press, New Jersey
- Damaschun, G, Damaschun, H, Fabian, H, Gast, K, Kröber, R, Wieske, M & Zirwer, D 2000, '*Conversion of Yeast Phosphoglycerate Kinase Into Amyloid-Like Structure*', *PROTEINS: Structure, Function, and Genetics*, vol. 39, pp. 204 – 211.
- D. E. Shaw Research 2012, '*Desmond Tutorial*', viewed 16 September 2013, http://www.deshawresearch.com/Desmond_Tutorial_0.6.1.pdf
- Desmond Molecular Dynamics System, version 3.1, D. E. Shaw Research, New York, NY, 2012. Maestro-Desmond Interoperability Tools, version 3.1, Schrödinger, New York, NY, 2012.
- Dingermann, T, Hänsel, R & Zündorf, I, 2002, '*Pharmazeutische Biologie*', Springer, Berlin Heidelberg New York
- Drake, GWF 1996, *Atomic, Molecular, & Optical Physics Handbook*, American Institute of Physics, Woodbury NY
- Fasman, GD 1996, *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York

- Fermentas, Life Sciences 2006, *GeneJET™ Plasmid Miniprep Kit: Fast isolation of high quality plasmid DNA*, viewed 5 February 2013, http://www.cosmobio.co.jp/product/products_FER_20060307/GeneJET_plasmidMiniprepKit1.pdf
- Förster, Th 1948, 'Zwischenmolekulare Energiewanderung und Fluoreszenz', *Annalen der Physik*, 6. Folge, Band 2, pp. 55-75.
- Galisteo, ML, Mateo, PL & Sanchez-Ruiz, JM 1991, 'Kinetic Study on the Irreversible Thermal Denaturation of Yeast Phosphoglycerate Kinase', *Biochemistry*, vol. 30, pp. 2061-2066.
- Galla, HJ 1988, *Spektroskopische Methoden in der Biochemie*, Georg Thieme Verlag, Stuttgart
- GE Healthcare n.d.b, *Ion Exchange Chromatography - Instructions 71-5009-64 AE*, viewed 01 April 2013, https://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314742967685/litdoc71500964AE_20110831002945.pdf
- Ghosal, S & Srivastava, AK 2009, 'Fundamentals of Bioanalytical Techniques and Instrumentation', Raj Press, New Delhi
- Gill, SC & Hippel, PH 1989, 'Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data', *Analytical Biochemistry*, vol.182, pp. 319-326.
- Greenfield, NJ 2006, 'Using circular dichroism spectra to estimate protein secondary structure', *Nature Protocols*, vol-1, pp. 2876-2890.
- Guex, N. and Peitsch, M.C. (1997) **SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling.** *Electrophoresis* **18**, 2714-2723.
- Hartl, FU, Bracher, A & Hayer-Hartl, M 2011, 'Molecular chaperones in protein folding and proteostasis', *Nature*, vol. 475, pp. 324-332
- The Hebrew University of Jerusalem n.d., 'Bacterial Glycerol Stocks', viewed 14 February 2013, http://wolfson.huji.ac.il/expression/procedures/bacterial/glycerol_stocks.html
- Herman, B, Frohlich, VEC, Lakowicz, JR, Fellers, TJ & Davidson, MW 2012, 'Fluorescence Resonance Energy Transfer (FRET) Microscopy', viewed 23 July 2013, <http://www.olympusmicro.com/primer/techniques/fluorescence/fret/fretintro.html>
- Hermanson, GT 2008, 'Bioconjugate Techniques', 2nd edn., Academic Press, London
- Holtzhauer, M 1997, 'Biochemische Labormethoden', 3rd edn., Springer Labor Manual, Berlin

- Holtzhauer, M 2006, '*Basic Methods for the Biochemical Lab*', Springer, Berlin
- Hornak, JP 1999, '*Molecular Fluorescence Spectroscopy*', viewed 09 July 2013, <http://jpkc.henu.edu.cn/sybpfx/scha/fluor.htm>
- Hug, H & Reiser, W 2000, '*Physikalische Chemie*', Europa-Lehrmittel, Haan-Gruiten
- Jorgensen, WL, Maxwell, DS & Tirado-Rives, J 1996, '*Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids*', J. Am. Chem. Soc., vol. 118, pp. 11225-11236.
- Lajtha, A & Banik, N 2007, '*Handbook of Neurochemistry and Molecular Neurobiology*', 3rd edn., Springer, New York
- Lakowicz, JR 2006, '*Principles of Fluorescence Spectroscopy*', 3rd edn., Springer, New York
- Lemkul, J 2012, '*GROMACS Tutorial – Lysozyme in Water*', Department of Biochemistry, Virginia Tech, viewed 16 September 2013, <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/>
- Life Technology Corporation 2013a, '*Fluorescence SpectraViewer*', viewed 07 July 2013, <http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cell-Analysis/Labeling-Chemistry/Fluorescence-SpectraViewer.html>
- Life Technology Corporation 2013b, '*Fluorescence quantum yields (QY) and lifetimes (τ) for Alexa Fluor dyes—Table 1.5*', viewed 08 October 2013, <http://www.lifetechnologies.com/at/en/home/references/molecular-probes-the-handbook/tables/fluorescence-quantum-yields-and-lifetimes-for-alexa-fluor-dyes.html>
- Lillo, MP, Szpikowska, BK, Mas, MT, Sutin, JD & Beechem, JM 1997, '*Real-Time Measurement of Multiple Intramolecular Distances during Protein Folding Reactions: A Multisite Stopped-Flow Fluorescence Energy-Transfer Study of Yeast Phosphoglycerate Kinase*', Biochemistry, vol. 36, pp. 11273-11281.
- Löffler, G, Petrides, PE & Heinrich, PC 2007, '*Biochemie & Pathobiochemie*', 8th edn., Springer, Heidelberg
- Maestro, version 9.3, Schrödinger, LLC, New York, NY, 2012.
- Merck Millipore 2013, '*pET-15b Vector TB045 5/99*', viewed 13 February 2013, http://www.merckmillipore.de/life-science-research/vector-table-novagen-pet-vector-table/c_HdSb.s1O77QAAAEhPqsLdcab?PortalCatalogID=merck4biosciences&CountryName=Austria

- MicroCal n.d., *VP-DSC MicroCalorimeter User's Manual – MAU120030 Rev B*, viewed 16 May 2013, http://www.uic.edu/orgs/ctrstbio/manuals/vpdsc_manual.pdf
- MicroCal 1998, *DSC Data Analysis in Origin® - Tutorial Guide*, viewed 22 May 2013, <http://www.uic.edu/orgs/ctrstbio/manuals/dscOrigin.pdf>
- Modeller, version 9.10, Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, California Institute for Quantitative Biomedical Research, San Francisco, 2012
- MOE (Molecular Operating Environment), version 2008.10; software available from Chemical Computing Group Inc. 1010 Sherbrooke Street West, Suite 910, Montreal, Quebec, Canada H3A 2R7, **2008**: <http://www.chemcomp.com>
- Müller-Esterl, W 2004, '*Biochemie Eine Einführung für Mediziner und Naturwissenschaftler*', Spektrum, München
- Nickson, AA & Clarke, J 2010, 'What lessons can be learned from studying the folding of homologous proteins?', *Methods*, vol. 52, pp. 38-50
- Nobel Media AB 2013, '*The Nobel Prize in Physiology or Medicine 2013*', viewed 4 November 2013, http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/
- Osváth, S, Köhler, G, Závodszky & Fidy, J 2005, '*Asymmetric effect of domain interactions on the kinetics of folding in yeast phosphoglycerate kinase*', *Protein Science*, vol. 14, pp. 1609-1616.
- Richter, G 2003, *Praktische Biochemie: Grundlagen und Techniken*, Thieme, Stuttgart
- Roe, S 2001, *Protein Purification Techniques*, 2nd edn, Oxford University Press, Oxford
- Rosenkranz, T 2011a, '*Time-Resolved Single Molecule FRET Studies on Folding/Unfolding Transitions and on Functional Conformational Changes of Phosphoglycerate Kinase*', Forschungszentrum Jülich, Jülich
- Rothman, RH 2012, '*Agarose Gel Electrophoresis*', viewed 14 February 2013, <http://people.rit.edu/rhrsbi/GEPages/LabManualPDF5ed/15%20electrophoresis.pdf>
- Sali, A 2012, '*Tutorial Advanced example: Modeling of a protein-ligand complex based on multiple templates, loop refinement and user specified restraints.*', viewed 2 November 2012, <http://salilab.org/modeller/tutorial/advanced.html>
- Schlegel, HG 1992, '*Allgemeine Mikrobiologie*', 7th edn., Thieme, New York

- Setlow, JK 2007, '*Genetic Engineering Principles and Methods*', vol. 28, Springer, New York
- Sivasankar, B 2005, '*Bioseparations: Principles and Techniques*', Prentice Hall of India Private Limited, New Delhi
- Søndergaard, CR, Olsson, MHM, Rostkowski, M and Jensen, JH 2011, "*Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values*" *Journal of Chemical Theory and Computation*, vol. 7, pp. 2284-2295.
- Spoel, D, Lindahl, E, Hess, B, Buuren, AR, Apol, E, Meulenhoff, PJ, Tieleman, DP, A. L. T. M. Sijbers, ALTM, K. A. Feenstra, KA, R. Van Drunen, R and H. J. C. Berendsen, HJC 2010, '*Gromacs User Manual version 4.5.6*', viewed 15 September 2013, <ftp://ftp.gromacs.org/pub/manual/manual-4.5.6.pdf>
- Stanford University 2003, *Transformation Protocol Using Heat Shock*, viewed 3 February 2013, <http://www.stanford.edu/~teruel1/Protocols/pdf/Transformation%20Protocol%20Using%20Heat%20Shock.pdf>
- Strucksberg, KH, Rosenkranz, T & Fitter, J 2007, '*Reversible and irreversible unfolding of multi-domain proteins*', *Biochimica et Biophysica Acta*, vol. 1774, pp. 1591-1603.
- Thermo Scientific 2011, *TECH TIP #31: Calculate dye:protein (F/P) molar ratio*, viewed 10 April 2013, <http://www.piercenet.com/files/TR0031-Calc-FP-ratios.pdf>
- Thermo Scientific 2013a, *GeneRuler 1 kb DNA Ladder 250 to 10,000 bp*, viewed 13 February 2013, <http://www.thermoscientificbio.com/nucleic-acid-electrophoresis/generuler-1-kb-dna-ladder-250-to-10000-bp/>
- Thermo Scientific 2013b, *PageRuler Plus Prestained Protein Ladder*, viewed 26 February 2013, <http://www.thermoscientificbio.com/protein-electrophoresis/pageruler-plus-prestained-protein-ladder/?rdr=true&terms=Pageruler>
- Wegerhoff, R, Weidlich, O & Kässens, M 2007, '*Basics of Light Microscopy & Imaging*', Olympus, GIT Verlag, Darmstadt
- Xiong, J 2006, '*Essential bioinformatics*', Cambridge University Press, New York
- Yon-Kahn, J & Hervé, G 2010, '*Molecular and Cellular Enzymology*', vol. II, Springer, Berlin
- Ziegenhorn, J, Seen, M, Bücher, T 1976, '*Molar Absorptivities of β -NADH and β -NADPH*', *CLIN. CHEM.*, vol. 22, no. 2, pp. 151-160.

9 Declaration

Hereby, I declare that this work was only written by me and that I have not used any prohibited aids. Furthermore, I confirm that all sources are marked by a quote whether they were used directly or indirectly. I have also tried to contact every owner of copyright protected images and obtained their agreement to use those pictures for this work. If any irregularities within this work appear according to copyright violations, I would like to be informed. Moreover, I certify that this version of the work has not been submitted or published before.

10 Appendix

10.1.1 myPGK sequence

Sequence of the modified yeast phosphoglycerate kinase according to the dissertation of Dr. Tobias Rosenkranz (2011a, p.139):

yPGK C97S Q135C S290C = myPGK

MGSSHHHHHH SSGLVPRGS M¹SLSSKLSVQ DLDLKD KRVF IRVDFNVPLD
GKKITSNQRI VAALPTIKYV LEHHPRYVVL ASHLGRPNGE RNEKYS LAPV
AKELQSLLGK DVTFLNDSVG PEVEAAVKAS APGSVILLEN LRYHIEEEGS
RKVDGCKVKA SKEDVQKFRH ELSSLADVYI NDAFGTAHR AHSSMVGF DL
PQRAAGF LLE KELKYFGKAL ENPTRPFLAI LGGAKVADKI QLIDNLLDKV
DSIIIGGGMA FTFKKVLENT EIGDSIFDKA GAEIVPKLME KAKAKGVEVV
LPVDFIADA FCADANTKT V TDKEGIPAGW QGLDNGPESR KLFAATVAKA
KTIVWNGPPG VFEFEKFAAG TKALLDEVVK SSAAGNTVII GGGDTATVAK
KYGVTDKISH VSTGGGASLE LLEGKELPGV AFLSEKK **PAANKARKEA**
ELAAATAEQ

10.1.2 Modeller – Homology modelling alignment

aln.pos	10	20	30	40	50	60	70	
1QPGA	-----	SLSSKLSVQDLDLKD KRVF IRVDFNVPLD	GKKITSNQRI	-----				
2K4ZA	-----							
1DVJA	MDVM-----	NRLILAM-----			DLMN---	RDDALRVTGEVREYIDTVKIGYPLVLS		
myPGK	MGSSHHHHHHSSGLVPRGSM	SLSSKLSVQDLDLKD KRVF IRVDFNVPLD	GKKITSNQRI	-----	V-----			
aln.pos	80	90	100	110	120	130	140	150
1QPGA	-----	AALPTIKYVLEHHPRYVVLASHLGQPN	GERNEKYS LAPVAKELQSLLGKDVTFLNDCV	GPEVEAAVKA				
2K4ZA	-----							
1DVJA	EGMDIIAEFRKRFGC-----							
myPGK	-----	A--ALPTIKYVLEHHPRYVVLASHLGRP	NGERNEKYS LAPVAKELQSLLGKDVTFLNDSV	GPEVEAAVKA				
aln.pos	160	170	180	190	200	210	220	230
1QPGA	SAPGSVILLENLRYHIEEEGSRKVDGCKVK	ASK-EDVQKFRHELSSLADVYINDAFGTAH	-----	RAHSSMV-				
2K4ZA	-----							
1DVJA	-----					RIIADFKVADIPETNEKICRATFKA		
myPGK	SAPGSVILLENLRYHIEEEGSRKVDGCKVKA	ySKEDVQKFRHELSSLADVYINDAFGTAH	-----	RAHSSMVG				
aln.pos	240	250	260	270	280	290	300	310
1QPGA	-----	GFDLPQRA----	AGFLLEKELKYFGKALENP-----					
2K4ZA	-----							
1DVJA	GADAIIVHGFP	GADSVRACLNVAEE-----	MGRE-----					

10 Appendix

```
myPGK      FDLP-----Q-----RAAGFLEKE-----LKYFGKALENP-----

aln.pos      320      330      340      350      360      370      380      390
1QPGA      --TRPFLAI-----LGGAKV--ADKIQLIDNLLDK-----VDSIIIGGGMAFTFKKVLENTEIGD
2K4ZA      GTEGMCLRL-----AAGR-----N
1DVJA      -----VFLITEMSHPGAEMFIQGADEIARMGVDLGVKNYVGPSTRPERLSRL-----
myPGK      --TRPFLAI-----LGGAKV--ADKIQLIDNLLDK-----VDSIIIGGGMAFTFKKVLENTEIGD

aln.pos      400      410      420      430      440      450      460
1QPGA      SIFDKAGAEIVPKLMEKAKAGVEVVLVPVDFIIADAFSADANTKTVTDKEGIPAGWQGLDNGPESRKLFAATVAKA--
2K4ZA      PD-----GSIDYR-----MGFDDL
1DVJA      -----REI
myPGK      SIFDKAGAEIVPKLMEKAKAGVEVVLVPVDFIIADAFCADANTKTVTDKEGIPAGWQGLDNGPESRKLFAATVAKA--

aln.pos      470      480      490      500      510      520      530      540
1QPGA      -----KTIVWN-----GP-----PGVFEFEKFAAGTKALLDEVVKSSAAGNTVIIGG
2K4ZA      TEDDIRLTSEGVEIVIAPDYVSLDQTTLDYVELEPGQFHFIFLNPRDPTY-----
1DVJA      IG-----QDSFLIS-----PGV-----GAQGG-----DPGETLRF-----ADAI--
myPGK      -----KTIVWN-----GP-----PGVFEFEKFAAGTKALLDEVVKSSAAGNTVIIGG

aln.pos      550      560      570      580      590      600      610      620
1QPGA      GDTATVA-----KKYGVTDKISHVSTGGGASLELLEGKELPGVAFLEKK-----
2K4ZA      -----RPPSGG-----
1DVJA      --VGRSIYLDADNPAAAAAGIIESI-----KDLLIPEDPAANKARKEAE
myPGK      GDTATVA-----KKYGVTDKISHVSTGGGASLELLEGKELPGVAFLEKKPAANKARKEAEAAATAEQ-----

aln.pos      630
1QPGA      -----
2K4ZA      -----
1DVJA      LAAATA
myPGK      -----
```

10.1.3 GROMACS – setup

Addition of ions script (Lemkul 2012):

```
; ions.mdp - used as input into grompp to generate ions.tpr
; Parameters describing what to do, when to stop and what to save
integrator      = steep          ; Algorithm (steep = steepest descent
minimization)
emtol           = 1000.0          ; Stop minimization when the maximum
force < 1000.0 kJ/mol/nm
emstep         = 0.01            ; Energy step size
nsteps         = 50000           ; Maximum number of (minimization) steps
to perform

; Parameters describing how to find the neighbors of each atom and how
to calculate the interactions
nstlist        = 1              ; Frequency to update the neighbor list
and long range forces
ns_type        = grid           ; Method to determine neighbor list
(simple, grid)
rlist          = 1.0            ; Cut-off for making neighbor list
(short range forces)
```

```

coulombtype      = PME                ; Treatment of long range electrostatic
interactions
rcoulomb         = 1.0                ; Short-range electrostatic cut-off
rvdw            = 1.0                ; Short-range Van der Waals cut-off
pbc              = xyz                ; Periodic Boundary Conditions (yes/no)

```

Energy minimization script (Lemkul 2012):

```

; minim.mdp - used as input into grompp to generate em.tpr
; Parameters describing what to do, when to stop and what to save
integrator       = steep              ; Algorithm (steep = steepest descent
minimization)
emtol            = 1000.0              ; Stop minimization when the maximum
force < 1000.0 kJ/mol/nm
emstep           = 0.01               ; Energy step size
nsteps           = 50000              ; Maximum number of (minimization) steps
to perform

; Parameters describing how to find the neighbors of each atom and how
to calculate the interactions
nstlist          = 1                  ; Frequency to update the neighbor list
and long range forces
ns_type          = grid               ; Method to determine neighbor list
(simple, grid)
rlist            = 1.0                ; Cut-off for making neighbor list
(short range forces)
coulombtype      = PME                ; Treatment of long range electrostatic
interactions
rcoulomb         = 1.0                ; Short-range electrostatic cut-off
rvdw            = 1.0                ; Short-range Van der Waals cut-off
pbc              = xyz                ; Periodic Boundary Conditions (yes/no)

```

NVT equilibration script (Lemkul 2012):

```

title           = OPLS Lysozyme NVT equilibration
define          = -DPOSRES           ; position restrain the protein
; Run parameters
integrator       = md                 ; leap-frog integrator
nsteps           = 50000              ; 2 * 50000 = 100 ps
dt              = 0.002              ; 2 fs
; Output control
nstxout          = 100                ; save coordinates every 0.2 ps
nstvout          = 100                ; save velocities every 0.2 ps
nstenergy        = 100                ; save energies every 0.2 ps
nstlog           = 100                ; update log file every 0.2 ps
; Bond parameters
continuation     = no                 ; first dynamics run
constraint_algorithm = lincs          ; holonomic constraints
constraints      = all-bonds          ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter       = 1                  ; accuracy of LINCS

```

```
lincs_order      = 4                ; also related to accuracy
; Neighborsearching
ns_type          = grid             ; search neighboring grid cells
nstlist          = 5                ; 10 fs
rlist            = 1.0              ; short-range neighborlist cutoff (in
nm)
rcoulomb         = 1.0              ; short-range electrostatic cutoff (in
nm)
rvdw             = 1.0              ; short-range van der Waals cutoff (in
nm)
; Electrostatics
coulombtype      = PME              ; Particle Mesh Ewald for long-range
electrostatics
pme_order        = 4                ; cubic interpolation
fourierspacing   = 0.16            ; grid spacing for FFT
; Temperature coupling is on
tcoupl          = V-rescale         ; modified Berendsen thermostat
tc-grps         = Protein Non-Protein ; two coupling groups - more
accurate
tau_t           = 0.1  0.1         ; time constant, in ps
ref_t           = 300  300         ; reference temperature, one for each
group, in K
; Pressure coupling is off
pcoupl          = no               ; no pressure coupling in NVT
; Periodic boundary conditions
pbc             = xyz             ; 3-D PBC
; Dispersion correction
DispCorr        = EnerPres        ; account for cut-off vdW scheme
; Velocity generation
gen_vel         = yes             ; assign velocities from Maxwell
distribution
gen_temp        = 300             ; temperature for Maxwell distribution
gen_seed        = -1             ; generate a random seed
```

NPT equilibration script (Lemkul 2012):

```
title           = OPLS Lysozyme NPT equilibration
define          = -DPOSRES         ; position restrain the protein
; Run parameters
integrator      = md               ; leap-frog integrator
nsteps          = 50000            ; 2 * 50000 = 100 ps
dt              = 0.002            ; 2 fs
; Output control
nstxout         = 100              ; save coordinates every 0.2 ps
nstvout         = 100              ; save velocities every 0.2 ps
nstenergy       = 100              ; save energies every 0.2 ps
nstlog          = 100              ; update log file every 0.2 ps
; Bond parameters
continuation    = yes              ; Restarting after NVT
constraint_algorithm = lincs        ; holonomic constraints
constraints     = all-bonds        ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter      = 1                ; accuracy of LINCS
lincs_order     = 4                ; also related to accuracy
; Neighborsearching
```

```

ns_type           = grid           ; search neighboring grid cells
nstlist           = 5              ; 10 fs
rlist             = 1.0            ; short-range neighborlist cutoff (in
nm)
rcoulomb          = 1.0            ; short-range electrostatic cutoff (in
nm)
rvdw              = 1.0            ; short-range van der Waals cutoff (in
nm)
; Electrostatics
coulombtype       = PME            ; Particle Mesh Ewald for long-range
electrostatics
pme_order         = 4              ; cubic interpolation
fourierspacing    = 0.16          ; grid spacing for FFT
; Temperature coupling is on
tcoupl            = V-rescale      ; modified Berendsen thermostat
tc-grps           = Protein Non-Protein ; two coupling groups - more
accurate
tau_t             = 0.1 0.1        ; time constant, in ps
ref_t             = 300 300        ; reference temperature, one for each
group, in K
; Pressure coupling is on
pcoupl            = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype        = isotropic      ; uniform scaling of box vectors
tau_p             = 2.0            ; time constant, in ps
ref_p             = 1.0            ; reference pressure, in bar
compressibility    = 4.5e-5         ; isothermal compressibility of water,
bar^-1
refcoord_scaling  = com
; Periodic boundary conditions
pbc               = xyz            ; 3-D PBC
; Dispersion correction
DispCorr          = EnerPres       ; account for cut-off vdW scheme
; Velocity generation
gen_vel           = no             ; Velocity generation is off

```

Molecular dynamic simulation script (Lemkul 2012):

```

title            = OPLS Lysozyme MD
; Run parameters
integrator        = md             ; leap-frog integrator
nsteps           = 500000          ; 2 * 500000 = 1000 ps, 1 ns
dt               = 0.002           ; 2 fs
; Output control
nstxout          = 1000            ; save coordinates every 2 ps
nstvout          = 1000            ; save velocities every 2 ps
nstxtcout        = 1000            ; xtc compressed trajectory output every
2 ps
nstenergy        = 1000            ; save energies every 2 ps
nstlog           = 1000            ; update log file every 2 ps
; Bond parameters
continuation      = yes            ; Restarting after NPT
constraint_algorithm = lincs       ; holonomic constraints
constraints       = all-bonds      ; all bonds (even heavy atom-H bonds)
constrained
lincs_iter       = 1               ; accuracy of LINCS

```

```
lincs_order      = 4                ; also related to accuracy
; Neighborsearching
ns_type          = grid             ; search neighboring grid cells
nstlist          = 5                ; 10 fs
rlist            = 1.0              ; short-range neighborlist cutoff (in
nm)
rcoulomb         = 1.0              ; short-range electrostatic cutoff (in
nm)
rvdw             = 1.0              ; short-range van der Waals cutoff (in
nm)
; Electrostatics
coulombtype      = PME              ; Particle Mesh Ewald for long-range
electrostatics
pme_order        = 4                ; cubic interpolation
fourierspacing   = 0.16            ; grid spacing for FFT
; Temperature coupling is on
tcoupl          = V-rescale         ; modified Berendsen thermostat
tc-grps         = Protein Non-Protein ; two coupling groups - more
accurate
tau_t            = 0.1  0.1         ; time constant, in ps
ref_t            = 300  300         ; reference temperature, one for each
group, in K
; Pressure coupling is on
pcoupl          = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype      = isotropic         ; uniform scaling of box vectors
tau_p           = 2.0              ; time constant, in ps
ref_p           = 1.0              ; reference pressure, in bar
compressibility  = 4.5e-5           ; isothermal compressibility of water,
bar^-1
; Periodic boundary conditions
pbc             = xyz               ; 3-D PBC
; Dispersion correction
DispCorr        = EnerPres         ; account for cut-off vdW scheme
; Velocity generation
gen_vel         = no               ; Velocity generation is off
```

11 Abstract

The aim of this thesis was to characterise the thermal unfolding profile of the C97S Q135C S290C phosphoglycerate kinase mutant from *Saccharomyces cerevisiae* between 15 °C and 75 °C. This effort was accomplished by utilizing experiments as well as simulations. The CD measurements revealed a melting temperature at 39,9 °C where half of the secondary structures were vanished. All detectable structures were gone over 47 °C and new ones appeared over 56 °C to due protein aggregation. The effect of aggregation was also seen during the DSC upscan and it depended strongly on the protein concentration. Furthermore, the DSC results showed an melting temperature at 53,16 °C where half of the protein was unfolded. Moreover, the thermodynamic results proved that the unfolding process is more complicated than a simple two-state transition. In addition, information about the temperature dependent interdomain distance profile was obtained by measuring the distance between CYS155 and CYS310 via FRET and complemented by the outcome of MD simulations. The results showed that the distance decreased until 42 °C to an average of 40 Å. Then, the domains remained mainly at a distance of 48 Å between 46 °C and 54 °C before they moved about 12 Å further apart at higher temperatures due to protein unfolding. As a consequence of these results the unfolding profile of the modified yeast phosphoglycerate kinase can be summarized in the following scheme: $N \rightarrow I \rightarrow U \rightarrow F$. This folding path illustrates the transition of the protein from the native state N to an intermediate state I which is characterized by an interdomain distance of 48 Å and the lack of detectable secondary structures. Furthermore, the protein reaches the unfolded state U at higher temperatures before it forms irreversible aggregates in a final state F.

12 German abstract - Zusammenfassung

Zielsetzung dieser Diplomarbeit war die Analyse der thermischen Entfaltung einer Phosphoglyceratkinase Mutante (C97S Q135C S290C) von *Saccharomyces cerevisiae* zwischen 15 °C und 75 °C. Im Rahmen dessen wurden Experimente und Simulationen durchgeführt. CD Messungen ergaben, dass das Protein einen Schmelzpunkt von 39,9 °C hat. Dieser Punkt ist dadurch charakterisiert, dass sich die Hälfte aller Sekundärstrukturen aufgelöst hat. Über 47 °C waren alle Strukturen verschwunden, bevor neue Strukturen über 56 °C aufgrund von Proteinaggregation nachweisbar wurden. Die Effekte der Aggregation wurden ebenfalls bei den DSC upscan Messungen detektiert und waren stark von der Proteinkonzentration abhängig. Weiters zeigten die DSC Ergebnisse einen Schmelzpunkt bei 53,16 °C, welcher jenen Punkt markiert, an dem die Hälfte der Proteine denaturiert war. Darüber hinaus belegten die extrahierten thermodynamischen Parameter, dass die Entfaltung komplexer ist und nicht monophasisch abläuft. Daneben wurden weitere Informationen über die temperaturabhängige Abstandsverteilung zwischen den Domänen durch FRET Messungen ermittelt. Dazu wurde der Abstand zwischen CYS155 und CYS310 gemessen und mit Ergebnissen von MD Simulationen korrigiert und verglichen. Die Profile zeigten, dass sich der Abstand verkleinert, bis dieser bei 42 °C einen durchschnittlichen Wert von 40 Å erreicht hat. Danach verharrten die Domänen zwischen 46 °C und 54 °C hauptsächlich bei 48 Å, bevor sich diese aufgrund der Entfaltung bei höheren Temperaturen um 12 Å weiter voneinander entfernten. Basierend auf diesen Resultaten lässt sich das Entfaltungsprofil von der PGK Mutante im folgenden Schema zusammenfassen: $N \rightarrow I \rightarrow U \rightarrow F$. Dieses Schema verdeutlicht, dass das Protein durch Temperaturerhöhung vom nativen Zustand N in einen intermediären Zustand I übergeht. Dieses Intermediat ist sowohl durch einen Abstand von 48 Å zwischen den Domänen als auch durch die Abwesenheit von messbaren Sekundärstrukturen gekennzeichnet. Darüber hinaus erreicht das Protein den ungefalteten Zustand U, bevor die Aggregation beginnt und das Protein im irreversiblen finalen Zustand F verbleibt.

13 Curriculum Vitae

Personal information:

Date of birth: June 21st 1985
Place of birth: Amstetten
Marital status: single
Nationality: Austria
Address: Untere Augartenstraße 31, 1020 Wien



Education:

10/2011 Diploma thesis: Department of Medicinal Chemistry, University of Vienna, in Cooperation with:
Theoretical Biochemistry Group, Institute for Theoretical Chemistry, University of Vienna
Biomolecular Optical Spectroscopy Group, Department of Structural and Computational Biology, University of Vienna
Institute of Structural Chemistry, Chemical Research Center, Hungarian Academy of Sciences
09/2011 2. Diploma examination
01/2006 1. Diploma examination
10/2004 - today Diploma study Pharmacy at the University of Vienna
June 21st 2004 Matura (a-level), average mark: „1,6“
09/1999 – 06/2004 Polytechnic Institute for Chemical Operating Technology (HTL) Wels
09/1995 – 06/1999 Secondary modern school for music: Blindenmarkt
09/1991 – 06/1995 Elementary school: St. Martin – Karlsbach

Work experience:

11/2011 Job at the “Laaerberg Apotheke”, Laaer Berg Straße 168, Wien
07/2008 & 07/2007 Internship at the “Schwan Apotheke”, Lessinggasse 23, Wien
07- 08/2002 Internship at the Austrian Federal Office for Safety and Healthcare GmbH Dairy Industry, Wolfpassing
07- 08/2001 Internship at the Federal Institute for Foodstuff Inspection, Linz